AVIAN HAEMOSPORIDIANS: DETECTION, HOST, AND CLIMATE ASSOCIATION ACROSS CONTRASTING REGIONS OF AFRICA

A Dissertation

by

JOHANNA ALEXANDRA HARVEY

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Gary Voelker
Committee Members,	Jessica Light
	Diana Outlaw
	Michel Slotman
Head of Department,	David Caldwell

May 2018

Major Subject: Wildlife and Fisheries Sciences

ABSTRACT

Avian haemosporidians make up one of the most widely distributed and diverse vector borne parasite systems, found nearly worldwide. While there has been growth in the use of avian haemosporidians as a model system for vector borne parasites, there remain unanswered questions in this system. Little remains known of the variation in results when assessing various source materials and how these recoveries bias our results. Further, a majority of avian haemosporidian diversity likely remains undiscovered particularly across sub-Saharan Africa. Each new recovery helps to further elucidate distributional patterns of diversification.

Few studies have addressed the importance of source material selection when assessing these relationships. We show that source material, here blood and pectoral muscle, do not yield equivalent results when assessing prevalence and genetic diversity of haemosporidian genera. We find higher prevalence and genetic diversity is recovered from blood versus pectoral muscle for *Haemoproteus*. Contrastingly, we find that a higher prevalence of *Plasmodium* is detected from pectoral muscle, while higher genetic diversity is recovered from blood. Results indicate that source material may bias parasite detection and be an important factor in study design.

We conducted the first sampling of avian haemosporidians, *Haemoproteus*, *Leucocytozoon*, and *Plasmodium* from Benin located in tropical West Africa and the Democratic Republic of the Congo. We sampled 222 and 421 birds, respectively, across distinct ecoregions with varied habitats. We detected haemosporidians in 113 of 222

individuals, resulting in a 50.9% infection rate in Benin. We detected haemosporidians in 187 of 421 individuals, resulting in a 42.4% infection rate in the Democratic Republic of the Congo. We molecularly recovered a high number of novel lineages, 52.9% and 71.4%, respectively. We characterized the multivariate variables which influence the distributions of haemosporidians genetic lineages, including host associations and bioclimatic variables for sampling from Benin. We introduce a novel visualization method to better capture the multivariate environment of haemosporidians, and this approach resulted in the recovery of intra-generic distribution patterns of diversity. We assessed host life history traits as a proxy for vector encounter rate and examined the variation in expressed traits across families with high sampling.

DEDICATION

This dissertation is dedicated to my family: William, Socorro, and Suzan.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Gary Voelker, and my committee members, Dr. Jessica Light, Dr. Diana Outlaw, and Dr. Michel Slotman, for their guidance and support throughout the course of this research.

I am grateful to the countries of Benin and the Democratic Republic of the Congo for the permission to perform this research. I would like to thank our Alphonse Adite and Robert Adite for the invaluable assistance in obtaining permits and field collecting in Benin. I would like to thank Drs. Gary Voelker, Toby J. Hibbitts and Jerry W. Huntley for the field collections of samples from Benin. I would like to thank Drs. Ben Marks, and Charles Kahindo for the field collections of samples from the Democratic Republic of the Congo.

I am grateful to Carter Atkinson for proposing valuable comments and suggestions that improved the chapter II.

I would like to thank Jon Puritz for being my constant support system and best champion. Thank you for being a sounding board for my ideas and analysis and giving advice. This would not have been possible without his support and constant encouragement. I would also like to thank Danielle Walkup for always being my friend and ally. Thank you for discussions and encouragement along the way and for being my writing and accountability buddy. I would also like to thank Jerry Huntley for being friend, advice giver, and support throughout this process. Undoubtedly, I would not have gotten here without him.

Finally, thanks to my mother and father for their unfailing encouragement throughout my long educational career.

CONTRIBUTORS AND FUNDING SOURCES

This dissertation was supervised by a doctoral committee consisting of Drs. Gary Voelker and Jessica Light of the Department of Wildlife and Fisheries Sciences. Dr. Voelker contributed intellectually to my research goals and questions. Dr. Jessica Light has over the years provided advice on methodology and interpretations of phylogenetic results. Outside committee member Dr. Michel Slotman from the department of Entomology has provided advice and input on methodology related to PCR and sequencing of parasites as well as advice to increase the rigorous determination of sequence data and biological units. Outside committee member Dr. Diana Outlaw from Mississippi State's Department of Biological Sciences has provided input on PCR, methodology and guidance in research questions.

Chapter II received an outside friendly review by Dr. Carter Atkinson from the USGS, Pacific Island Ecosystems Research Center.

Statistical advice on transformations was given by Dr. Wade Ryberg of the Natural Resources Institute.

All other work conducted for the dissertation was completed by the student independently.

Graduate study was supported by the Texas A&M Diversity fellowship. Post-candidacy support was received through the College of Agriculture and Life Science's Tom Slick Graduate Research Fellowship.

NOMENCLATURE

Bp	Base pairs
DRC	Democratic Republic of the Congo
Cyt b	Cytochrome oxidase beta
N	Number
mtDNA	Mitochondrial DNA
km	Kilometers
PCR	Polymerase chain reaction

TABLE OF CONTENTS

Page

Page

LIST OF FIGURES

Page

Page

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

Despite modern advances in medicine and research, disease outbreaks are still of great concern to humans and wildlife, and the majority of disease outbreaks are vector-mediated (Smith et al., 2014). Human malaria remains one of the leading causes of human morbidity, affecting nearly half of the human population (WHO, 2017). Malaria has been of biomedical interest for the better part of five thousand years. Chinese medical writings, Mesopotamian tablets from 2000 BC, and Egyptian papayri from 1570 BC describe the diagnostic symptoms, such as spleen enlargement and cyclical fevers, of a then unknown disease later credited to miasmas and later a bacterium (Garnham, 1966; Rich and Ayala, 2006; Cox, 2010). A better understanding of malaria began in the last century with Laveran's 1880 discovery of an exflagellating protozoa, part of the sexual phase that occurs with changes in temperature, as would occur when transferred from vertebrate host to vector, in a blood sample (Garnham, 1966; Cox, 2010). Nearly at the same time of Laveran's finding, Danilewsky detected a hematozoa infecting the blood of birds and reptiles, recognizing three genera of parasites: *Plasmodium*, *Haemoproteus*, and *Leucocytozoon* (Garnham, 1966). Danilewkys's work was further advanced by the discovery of methylene blue stain by Romanowsky (Garnham, 1966). Staining with methylene blue allowed the distinction of the parasite nucleus and cytoplasm from host cells, aiding in morphological identification and classification.

The group of parasites causing these diseases (malaria, haemoproteosis, or leucocytozoonosis) are unicellular apicomplexan protozoa that are digenetic (having obligate vertebrate and arthropod hosts) belonging to the family Haemosporidia. Parasites in this family are generally referred to as haemosporidians and infect terrestrial vertebrates (birds, mammals, and squamate reptiles) as their definitive hosts (Atkinson and Van Riper III, 1991; Valkiunas, 2005). While malaria parasites have in common basic morphology and life cycles (stages of sexual and asexual reproduction, i.e. life history traits), they differ widely in their dipteran and vertebrate host associations, location of merogony (erythrocitic or not), pigmentation, and shape, structure, and location of life stages within vertebrate hosts (Garnham, 1966; Atkinson and Van Riper III, 1991; Valkiunas, 2005).

The agents of human malaria include six *Plasmodium* species (Calderaro et al., 2013), and each cause devastating effects in humans. The well-studied human malaria species represent just a very small proportion of the diversity comprising of haemosporidians. The largest haemosporidian diversity has been recovered in birds (Bensch et al., 2009). Currently there are ~220 species morphologically described species of avian haemosporidians and 2,876 genetic lineages (MalAvi 2.3.3). Due to the advancement of molecular techniques for identification of haemosporidians, it is now thought that hematozoa diversity may be approximately equivalent or greater to host diversity (e.g., in Aves; Ricklefs and Fallon, 2002; Bensch et al., 2004). Despite the large number of genetic lineages recovered we still lack strong resolution of the avian haemosporidian phylogeny. This is due to a lack of host taxonomic sampling across

many bioregions and habitats (Clark et al., 2014; Outlaw et al., 2017). Sampling has been found to be biased toward temperate regions with much remaining to be discovered across tropical regions worldwide, but primarily with deficits in the Afrotropics and Asia (Clark et al., 2014). This discrepancy in sampling between temperate and tropical regions occurs despite the fact that avian species richness is highest in tropical regions (Mittermeier et al., 1998; Brooks et al., 2001).

Haemosporidian parasites can negatively impact avian populations, by affecting host fitness, from physical condition and reproductive success to reduced survival rates, as well as being a source of emerging disease in non-endemic ranges (Hamilton and Zuk, 1982; Merino et al., 2000; Marzal et al., 2005; Knowles et al., 2010; Asghar et al., 2011; Garcia-Longoria et al., 2014; Coon et al., 2016). This was most dramatically seen in the introduction of *P. relictum* to the Hawaiian Islands which resulted in extinctions and detrimental population declines forever altering the local avian fauna (van Riper III et al., 1986; Lapointe and Camp, 2013). Further, transmission can be facilitated by migration of birds, and this dispersal has been documented in malaria, West Nile Virus, and Lyme disease, among others (Ricklefs et al., 2005; Comstedt et al., 2006; Kilpatrick et al., 2006a; Garcia-Longoria et al., 2015). Recent findings in a New World study indicate that avian malaria speciation is primarily occurring via host switching (across bird species and taxonomic groups) in initial allopatry (i.e., geographic separation), as may occur during migration, dispersal events, and range shifts (Hellgren et al., 2007; Ricklefs et al., 2014).

Patterns of parasite distributions have been seen at localized landscape scales, but have not yet been discerned for haemosporidians at larger scales or across; this has impeded the identification of broad patterns biomes (Wood et al., 2007; Clark et al., 2014; Outlaw et al., 2017). Sub-Saharan African haemosporidian studies assessing broad avian communities have begun to recover a high diversity of novel lineages (Beadell et al., 2009; Okanga et al., 2014; Lutz et al., 2015). While studies focusing on a single or small target set of species have been informative for parasite biodiversity, they limit the overall recovered diversity, as host parasite specificity varies across lineages. Previous work has also been geographically limited, with taxonomically dense sampling only in Malawi, Nigeria, the western cape of South Africa, Gabon, and Cameroon (the former two restricted sampling to wetlands and marshes) (Beadell et al., 2009; Okanga et al., 2014; Lutz et al., 2015). Thus, large swaths of sub-Saharan Africa and its rich avifauna remain not sampled (Figure 1).

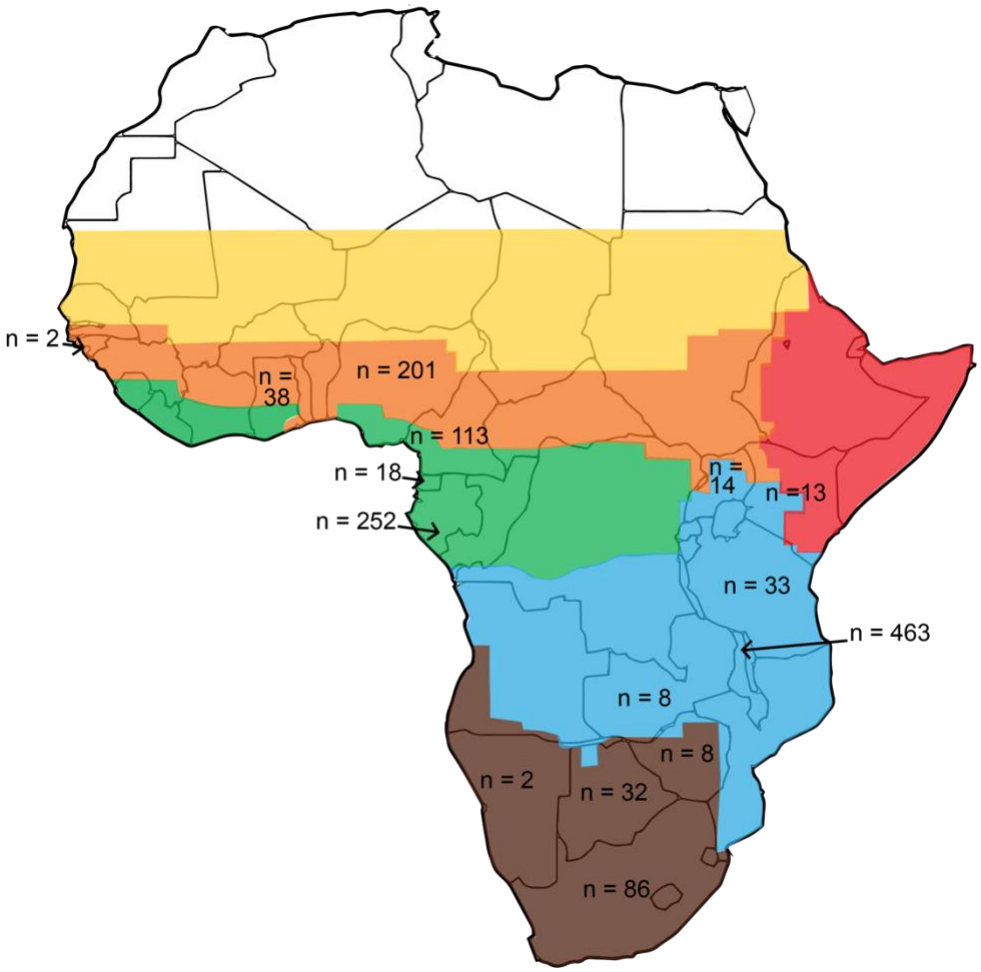

Figure 1. Map of avian bioregions with number (n) of currently recovered haemosporidian lineages recovered from sub-Saharan African countries (Adapted from Outlaw et al. 2017).

In this dissertation, I seek to fill knowledge gaps in several facets of haemosporidian research. Haemosporidian studies range in topics from host ecology and fitness effects to parasite evolution and haemosporidian diversity. Thus, the methods for molecular detection must be scrutinized to determine results across studies are equivalent. Here, I test the results across the most commonly collected source materials for avian haemosporidian detection studies. I increased sampling in sub-Saharan Africa

to two countries, Benin and the Democratic Republic of the Congo, with no previous sampling (Figure 1). Informing haemosporidian diversity with the detection of a large number of novel lineages and further informing distributional patterns with the recovery of previously detected lineages. This is my effort to improve the sampling and develop methods to better visualize the patterns of distribution in this highly diverse and complex parasite system.

CHAPTER II

AVIAN HAEMOSPORIDIAN DETECTION ACROSS SOURCE MATERIALS: PREVALENCE AND GENETIC DIVERSITY*

II.1 Introduction

Haemosporidian detection studies inform a broad field of research, from host ecology and fitness to parasite evolution and diversity, yet differences in results across source material have thus far been inconclusively assessed. Avian haemosporidian ranges nearly worldwide and consists of the genera *Haemoproteus*, *Leucocytozoon*, and *Plasmodium*. Infection resulting in malaria, haemoproteosis, or leucocytozoonosis (caused by *Plasmodium*, *Haemoproteus*, and *Leucocytozoon* infections, respectively) can affect host fitness, from individual physical condition and reproductive success to reduced survival rates, as well as exhibiting a selective force on populations (Hamilton and Zuk 1982; Merino et al. 2000; Marzal et al. 2005; Knowles et al. 2010; Asghar et al. 2011; García-Longoria et al. 2014; Coon et al. 2016). They can also have devastating population- and community-level effects in non-endemic ranges, as evidenced by avian responses to malarial introductions on the Hawaiian Islands (van Riper et al. 1986). Avian haemosporidian research has grown as a model study system, and as resulting

data has increased, we have begun to see the use of meta-studies to try and decipher broader evolutionary and ecological patterns (e.g., Clark et al. 2014; Outlaw et al. 2017).Overall, the majority of studies that have assessed avian haemosporidians have done so by targeting a small fragment of the mitochondrial DNA (mtDNA) cytochrome b (cyt b) gene (479 bps), the standard barcode, and the measure of lineage diversity. There are currently 2381 unique lineages documented in MalAvi, the avian haemosporidian database (MalAvi 2.2.9, December 17, 2016; Bensch et al. 2009), and descriptions of new variants occur regularly as the effects of malaria, haemoproteosis, and leucocytozoonosis on the ecology of their bird hosts are studied.

Avian haemosporidian parasite studies have predominantly sampled blood as the primary source material for parasite detection and analysis. Blood is most widely used due to its ease of collection, and the fact that birds are generally released unharmed, making it a more easily permitted and collected material. Previous work has found that *Leucocytozoon* detection did not vary when comparing blood collected from the brachial vein (circulation blood) to deep circulation blood (collected from the lungs), suggesting that location of blood draw will not influence detection (Holmstad et al. 2003). Further, various other tissue types (i.e., muscle, liver, and heart) are often used, as they may be widely accessible (i.e., already accessioned in natural history museum collections) or part of separate studies and collection efforts. Investigations most often utilize a single starting material due to study design, while other studies may use various source materials due to availability and logistics.

All haemosporidian genera undergo similar infection periods which include the following: prepatent, primary parasitemia (acute and chronic), latent stage, and secondary parasitemia (in response to relapse), within their vertebrate host. Vertebrate hosts, such as birds, are intermediate hosts and the location of asexual reproduction, while dipterans are the definitive host and location of sexual reproduction as well as the vector. However, there are distinct variations in life stages, patterns of development, and timing of transmission season across genera. All genera undergo development in both circulating blood and fixed tissues of hosts at various stages. Initial development for all genera occurs in fixed tissues (i.e., within parenchymal cells, spleen, liver, kidney heart, skeletal muscle, and endothelial lining of the capillaries) and is at the stage of trophozoites or meronts with the location of such being specific to genus and sometimes species (Valkiūnas 2005). At this stage (trophozoite or meront), either abortive development (i.e., dead end where parasite cannot develop into patent infection stages due to lack of host-parasite competence) occurs or the parasite continues to become a patent infection (Valkiūnas 2005). The patent circulating blood life stages are meronts and/or gametocytes depending on genera. Given these varying life-stages seen across haemosporidian genera, detection probability may be impacted across seasons and by host ecology (Mukhin et al. 2016).

Using single-species sampling, two previous studies have sought to address the differences in starting source material. Ramey et al. (2013) investigated *Haemoproteus*, *Leucocytozoon*, and *Plasmodium* in the Northern Pintail (*Anas acuta*) using paired blood and muscle source materials from 157 individuals. They found no significant difference

between sources and primarily recovered *Leucocytozoon*. However, they did find higher haplotype diversity for *Leucocytozoon* in blood (five additional haplotypes) versus muscle. Svennson-coehlo et al. (2016) screened 46 White-shouldered Fire-eyes (*Pyriglena leucoptera*) for *Haemoproteus* and *Plasmodium*, implementing a comprehensive-paired sampling regime across of source materials (including blood, brain, heart, pectoral muscle, and liver) using several screening methods. Their investigation only recovered *Plasmodium*, with no significant differentiation when comparing across all source materials, although detection rates did vary: e.g., blood prevalence (28%) was lower than pectoral muscle prevalence (36%), while heart had the highest prevalence (39%). Detection in combined tissues (liver, heart, and pectoral muscle) as compared to blood was significant.

Here, I investigate whether blood and pectoral muscle source materials result in equivalent detections of avian haemosporidian prevalence and genetic diversity. I examined a taxonomically diverse suite of avian hosts from the tropical West African country of Benin, with sampling encompassing varied habitat types. Assessing equivalency across source materials for detection of haemosporidians may inform biases in previous avian haemosporidian research. Broad comparisons of parasite detections are important for understanding the role parasites play in host ecology. Given that parasites play a powerful selective force on populations understanding how to best address distributions informs or ability to detect them and make meaningful comparison across studies.

II.2 Methods

II.2.1 Avian Sampling

Birds were sampled between May and June of 2010, as part of a broader systematics and biodiversity survey, at six sampling localities across Benin (Parc W: Point Triplo [11.89828N, 2.41149E]; Parc W: Chutes de Koudou [11.68006N, 2.31689E]; Lake Toho [7.54305N, 2.60687E]; Dogo Forest [7.54305N, 2.60687E]; Lama Forest [6.96026N, 2.16830E]; and Abomey Calavi [6.42292N, 2.34996E]) (Figure 2). These localities represent the diverse biogeographic regions and habitats found in Benin (see Linder et al. 2012). All localities have tropical savanna climates (Peel et al. 2007), with mean temperatures above 18 °C, a marked dry season and are differentiated by length and timing of wet seasons across Northern and Southern localities (Figure 2) Northern localities, ParcW localities are characterized by a long dry season with a pronounced single wet season, occurring from June to September (http:// worldclim.org; Hijmans et al. 2005). Sampling at these northern localities occurred between May 21st and 28th. All other southern localities (Dogo Forest, Lama Forest, Abomey Calavi, and Lake Toho) have two wet seasons, and sampling at these southern localities occurred from June 1st to 10th (http://worldclim.org; Hijmans et al. 2005).

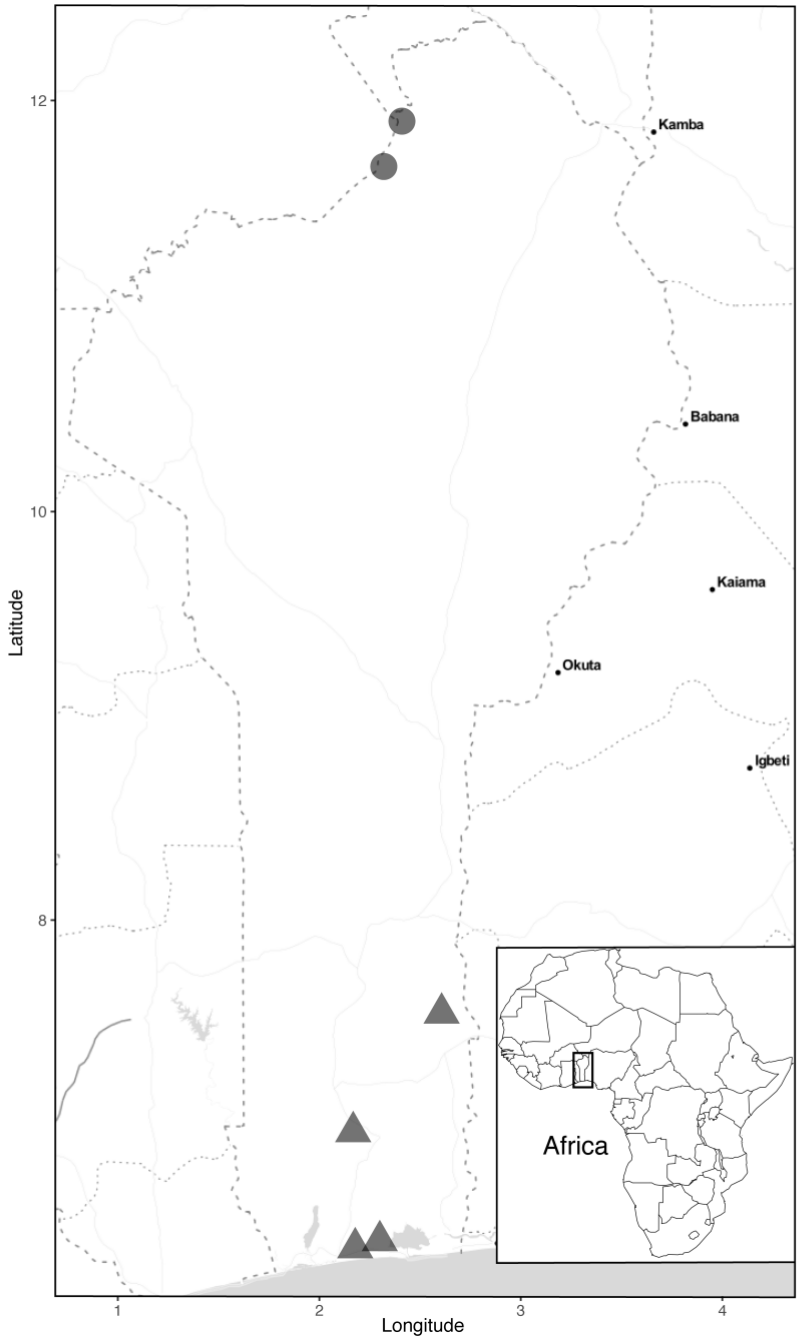

Figure 2. Map of Benin with northern sampling localities indicated by circles and southern localities indicated by triangles. Reprinted with permission from (Harvey et al. 2017).

Overall, 220 birds were captured via opportunistic mist-netting and sampled for blood and/or pectoral muscle tissue. Blood was sampled primarily directly from heart, after euthanasia. Blood was stored in Queen's Lysis buffer, and tissues were stored in 20% DMSO salt saturated storage buffer. All voucher specimens collected are accessioned in the Biodiversity and Research Teaching Collections, at Texas A&M. A combined total of 421 samples from 220 individuals (blood = 213; pectoral muscle = 208) were collected. For analysis, we used only paired blood and pectoral muscle samples (i.e., both taken from the same bird specimen), totaling 199 pairs. Avian sampling from the 199 paired sample individuals consisted of 7 orders, 29 families, and 78 species, with most sampling representing Passeriformes (80%). Paired samples (N = 199) represent 119 individuals from Northern localities and 80 individuals from Southern localities. Based on skull ossification data taken for each of the 199 individual hosts, our bird samples consisted of 157 adults and 23 juveniles, while ossification was not noted for 19 individuals. Individuals were not partitioned by age given the small number of juvenile individuals collected. All collecting of specimens was done under protocols approved by the Institutional Animal Care and Use Committee at Texas A&M University. Birds were sacrificed via cardio-thoracic compression, as suggested by the Ornithological Council's Guideline to the Use of Wild Birds in Research (Fair et al. 2010).

II.2.2 Molecular assessment of avian haemosporidians

We extracted whole genomic DNA from avian blood and pectoral muscle samples using the E.Z.N.A. Tissue extraction kit (Omega Bio-Tek, Norcross, GA) and

standard protocols except for the final elution was 75 µl to increase overall DNA yield. Polymerase chain reaction (PCR) was used to identify haemosporidian infection by amplifying a fragment of the mtDNA cyt b gene, and multiple primer pairs were used to amplify across known avian haemosporidian genetic diversity. Three primer pairs consisting of the same forward primer (UNIVF) and one of three distinct reverse primers UNIVR1, UNIVR2, or UNIVR3 were used, yielding 505, 535, and 571 base pairs (bp) (Drovetski et al. 2014), respectively. These primers encompass the entire 479 bp that are the standard region that is collected in the MalAvi avian haemosporidian database (Bensch et al. 2009). Collectively, these primers amplify all three genera of avian haemosporidians: *Haemoproteus* (to include subgenera Parahaemoproteus and *Haemoproteus*), *Leucocytozoon*, and *Plasmodium*. PCR amplification was carried out in 18.75 µl reaction containing 1× GoTaq Flexi buffer, 2.5 mM MgCl2, 0.2 mM of dNTP, 0.19 µl of each primer, and 0.9375 µ of GoTaq Flexi DNA polymerase (Promega Madison, Wisconsin, USA) with 1 µl of DNA template. Each sample (both blood and pectoral muscle) was tested via PCR for each primer pair, and up to three times per primer pair. If a positive amplicon was not detected and successfully sequenced during this process, it was considered the sample to negative for haemosporidians. All PCRs included four positive controls as well as two negative controls. Automated sequencing was performed bi-directionally using BigDye (Applied Biosystems, Foster City, CA, USA), and products were run out using an ABI 3730 at Beckman Coulter Genomics (Danvers, MA, USA).

Sequences were verified and aligned by eye using Geneious 6.1.8 (http://www.geneious.com;Kearseet al. 2012). Multiple infections were determined by the presence of double peaks on both chromatograms at one or more base positions. For sequences with double peaks, I PCR'd and re-sequenced two to three times to verify these double peaks. Sequences with < 3 double peaks were treated as single infections, due to a high probability; these are not multiple infections but instead the result of sequencing error (Szymanski and Lovette 2005). Verified multiple infections were further examined using the heterozygous plug-in in Geneious 6.1.8. Criteria to better qualify a multiple peak were > 70% peak similarity, a base calling confidence mean score of > 25 at each multiple peak site, and visual assessment of the strand quality. I implemented conservative criteria to remove the possibility of sequencing errors being erroneously called a multiple infection peak. Sequences with multiple peaks having met the above criteria (N =5) and were then assigned the appropriate IUPAC nucleotide ambiguity code for double infections.

All data were assessed phylogenetically and identified to genus by use of the MalAvi blast (Version 2.2.8, Bensch et al. 2009) and NCBI BLAST (Altschul et al. 1990) functions. Double infection data were phased using the Phase 2.1 (Stephens et al. 2001) implementation in DnaSP 5.10.1 (Librado and Rozas 2009) for reconstruction of single infection haplotypes. Phase input datasets were separated by genus and included all single infection data from our data along with top MalAvi blast (Version 2.2.8, Bensch et al. 2009) and NCBI BLAST matches. Phase was run for each genus with a 1000 iteration burn-in, 10 thinning intervals, and 1000 iterations. Phasing of data

reconstructs haplotypes from diploid gametic alleles and therefore may have flaws related to its ability to reliably assess multiple infections haplotypes. Phased data (N = 10) were then added to the final sequence dataset.

II.2.3 Statistical analyses

All statistical analyses were performed in R 3.3.2 (R Core Team 2016). Datasets were assessed for normality using Shapiro-Wilks test and assessed using quantile-quantile plots. Data were partitioned in five ways for further analysis. Partitioning allowed us to determine an overall infection rate as a percentage of host individuals detected positive (successful pcr detection and definitive sequencing), as well as to determine the overall detection rates for each avian haemosporidian genus. In dataset A, I assessed all host individuals (N = 199) as binary (detected versus not detected) for avian haemosporidian detection, regardless of source material. In dataset B, all host individuals were assessed separately for each tissue source (i.e., for statistical purposes, blood and pectoral muscle were treated as dependent variables), and all samples were coded as binary (detected versus not detected) for comparison. I used a McNemar's chi-squared test as implemented in the exact 2 × 2 package of R (Fay 2010) for paired sample data to detect equality of detection rates across starting source materials (blood versus pectoral muscle). In dataset C, all host individuals were treated as paired samples (blood versus pectoral muscle) and analyzed counts of the number of detections found in blood versus pectoral muscle, using a dependent t test to determine mean difference of blood versus pectoral muscle detections. I included count data for host individuals determined as positive for detection, because I found one to three distinct lineages of

avian haemosporidian per positive individual (given blood and pectoral muscle sources and three distinct primer sets, allowing a total of six possible detections per individual plus possible double infections). In dataset D, all host individuals (N =199) were assessed for each parasite genus separately, and then within genus by blood versus pectoral muscle. I then examined genus specific subsets using McNemar's chi-squared test and binary coding of detection. Each genus was examined independently, as detectability may vary between blood and pectoral muscle due to variations across genera life cycles. In dataset E, I examined all recovered avian haemosporidian lineages (E1) from positive samples (detected and sequenced, N = 155 recovered from 103 host individuals), allowing us to assess the lineages recovered and their source material while removing any bias from individuals not infected. I separated lineages by genus (*Haemoproteus* (N =55); *Leucocytozoon* (N =11); *Plasmodium* (N = 89)) and starting source material. I then utilized McNemar's chi-squared test to determine within genus equality of detection for blood and pectoral muscle source materials. I reduced the number of infections produced from double infections, all of which were in the genus *Haemoproteus* and counted these as single infections (E2) to correct for the possible error in PHASE assignment (reduced E2 *Haemoproteus* N = 50).

II.2.4 Genetic diversity analysis

We also assessed dataset E3 to examine genetic diversity recovered across blood and pectoral muscle starting materials. The final sequence dataset (same as dataset E1) was separated by genus (as above: *Haemoproteus* (N =55), *Leucocytozoon* (N=11), and *Plasmodium* (N= 89) and assessed for haplotype diversity using PopART 1.7

(http://popart.otago.ac.nz) implementing the minimum spanning for networks (Bandelt et al. 1999). Further, each sequence was coded for trait data as being recovered from blood, pectoral muscle, or both to determine differences in genetic diversity as resolved by each starting source material. Lineages recovered have been deposited in Genbank under accession numbers (MG018625 - MG018709) and the MalAvi database.

II.3 Results

II.3.1 Statistical analyses

For dataset A, detections at the individual host level (N = 199, blood and pectoral muscle combined) resulted in 103 individuals with one or more detections of avian haemosporidians, an overall 51.8% infection rate (*Haemoproteus* 22.6%, *Leucocytozoon* 6.5%, *Plasmodium* 34.2%) (Table 1).

For dataset B, 69 host individuals recovered detections in both blood and pectoral muscle material, 14 individuals had detections in blood alone, and 20 individuals had detections in pectoral muscle alone. Paired binary detections (detected/not detected) did not differ across starting source material (McNemar's X^2 (1) = 1.059, P = 0.303, Table 1, B).

Table 1. A) Detection rates for combined source materials. B) McNemar's test for binary detection of blood vs. pectoral muscle (muscle). C) Paired t-test using count detection data of blood vs. muscle. D) McNemar's test using binary detection assessing genera E1) McNemar's test for recovered lineages E2) Single infection assessment of genera. * denotes significant P values. Reprinted with permission from (Harvey et al. 2017)

A	Blood & Muscle	*N*	Detection Rate
	Overall	103	51.8%
	Haemoproteus	45	22.6%
	Leucocytozoon	13	6.5%
	Plasmodium	68	34.2%

B	Binary Detection	*N*	Blood Only	Muscle Only	Blood & Muscle	df	McNemar's $\chi 2$	*P*	Mean Blood (± SE)	Mean Muscle (± SE)
	Blood vs. Muscle	199	14	20	69	1	1.059	0.303	0.42 (0.035)	0.45 (0.035)

C	Count Detection	*N*	Blood Count	Muscle Count		df	Paired t-test	*P*	Mean Blood (± SE)	Mean Muscle (± SE)
	Blood vs. Muscle	199	105	105		198	0	1.0	0.528 (0.050)	0.528 (0.046)

D	Binary Detection	*N*	Blood Only	Muscle Only	Blood & Muscle	df	McNemar's $\chi 2$	*P*	Mean Blood (± SE)	Mean Muscle (± SE)
	Haemoproteus	199	11	7	20	1	0.889	0.346	0.226 (0.035)	0.176 (0.030)
	Leucocytozoon	199	4	1	8	1	1.8	0.18	0.060 (0.017)	0.045 (0.015)
	Plasmodium	199	12	25	26	1	4.568	0.032 *	0.247 (0.033)	0.323 (0.037)

E1	Recovered Lineages	*N*	Blood Only	Muscle Only	Blood & Muscle	df	McNemar's $\chi 2$	*P*	Mean Blood (± SE)	Mean Muscle (± SE)
	Haemoproteus	55	22	9	24	1	5.452	0.019 *	0.836 (0.050)	0.600 (0.067)
	Leucocytozoon	11	5	2	6	1	1.286	0.257	0.846 (0.104)	0.615 (0.140)
	Plasmodium	89	23	38	28	1	3.688	0.055	0.573 (0.053)	0.742 (0.047)

E2	E1- Only single infections	*N*	Blood Only	Muscle Only	Blood & Muscle	df	McNemar's $\chi 2$	*P*	Mean Blood (± SE)	Mean Muscle (± SE)
	Haemoproteus	50	18	9	24	1	3	0.083	0.824 (0.054)	0.647 (0.068)

E3	Haplotypes	*N*	Blood Only (%)	Muscle Only (%)	Blood & Muscle (%)
	Haemoproteus	37	13 (35.10%)	6 (16.22%)	18 (48.65%)
	Leucocytozoon	11	2 (18.18%)	3 (27.27%)	6 (54.54%)
	Plasmodium	33	11 (33.33%)	4 (12.12%)	18 (54.55%)

For dataset C, total counts of detections were equal (P = 1.0) in the dependent samples t- test for paired count data (N = 199), with blood and pectoral muscle each having 105 detections (both have M = 0.528, Table 1, C). For dataset D, there was some detection differentiation evident for source material when assessed by genus (N = 199, paired binary infection state, Table 1, D). *Haemoproteus* was detected in both blood and pectoral muscle for 20 host individuals, while 11 detections where found in blood alone, and 7 were detected in pectoral muscle alone. Of the 199 hosts with paired sampling, 161 individuals resulted in no *Haemoproteus* detections, and within *Haemoproteus*, there was no significant differentiation across source material (P =0.346; Table 1, D). I detected *Leucocytozoon* in both blood and pectoral muscle for eight host individuals, while four detections were observed in blood alone, and one detection was in pectoral muscle alone, and 186 host individuals recovered no *Leucocytozoon* detections. As in *Haemoproteus*, I found no significant difference in *Leucocytozoon* detections across source material (P = 0.18; Table 1, D). I detected *Plasmodium* in both blood and pectoral muscle for 26 host individuals, while 12 detections were found in blood alone, 25 detections were found in pectoral muscle alone, and 136 host individuals resulted in no *Plasmodium* detections. Unlike *Haemoproteus* and *Leucocytozoon*, I found significant inequality in detection between blood and pectoral muscle in *Plasmodium* (P = 0.032; Table 1).

For dataset E1, where recovered lineages (detected and sequenced) across host individuals and their source material were separated by genus, I recovered 157 sequences, with an average of 0.789 infections per individual (range of one to four

infections per individual). For *Haemoproteus* (N = 55), 22 infections were recovered from blood alone, while nine infections were recovered from pectoral muscle alone, and 24 identical lineages were recovered from both source materials. This resulted in a highly significant differentiation of detections dependent upon source material for *Haemoproteus* (P =0.019; Table 1, E1). However, if I removed the double infection data and treat those as single infections (E2), with 18 detections in blood only and 9 in pectoral muscle only, I still recover a marginally significant difference in detections for blood and pectoral muscle (P = 0.083; Table 1, E2). I found few unique *Leucocytozoon* detections (N =11, Table 1, E1) with five lineage detections found in blood alone and two detections found in pectoral muscle alone, along with six identical lineage detections recovered in the same individual from both source materials. Resulting differences were not significant (P = 0.257; Table 1, E1). Lastly, for *Plasmodium* (N = 89), I recovered 23 lineages in blood alone, 38 lineages were recovered in pectoral muscle alone, and 28 identical detections were recovered in blood and pectoral muscle. Detection differences were marginally significant (P = 0.055; Table 1, E1).

II.3.2 Genetic Diversity

For dataset E3, using a network analysis for each genus, I recovered a total of 37 unique haplotypes for *Haemoproteus* from 57 sequences mtDNA cyt b sequences, 18 of which were recovered in both blood and pectoral muscle host samples (Figure 3). However, 11 *Haemoproteus* haplotypes were unique to blood, and 4 were unique to pectoral muscle, resulting in a 33.3% loss of recovered genetic diversity if sampling pectoral muscle alone for *Haemoproteus*, and a 19.4% loss of genetic diversity if

sampling blood alone. For *Leucocytozoon*, I detected a total of 11 unique haplotypes (Figure 3) from 13 mtDNA cyt b sequences, where 6 haplotypes were recovered in blood and pectoral muscle paired samples. I found two haplotypes unique to blood and one unique to pectoral muscle, resulting in a 22.2% loss of genetic diversity if sampling pectoral muscle alone, and a 11.1% loss in diversity if sampling blood alone. Finally, I recovered 31 unique haplotypes for *Plasmodium* from 87 mtDNA cyt b sequences (Figure 3). While 18 were shared between blood and pectoral muscle, 10 were unique to blood comprising 32.2% of genetic diversity recovered, and only three were unique to pectoral muscle comprising 9.7% of the diversity. This resulted in a 32.2% loss in recovered genetic diversity if sampling pectoral muscle alone, and a 9.7% loss of recovered genetic diversity if sampling blood alone for *Plasmodium*.

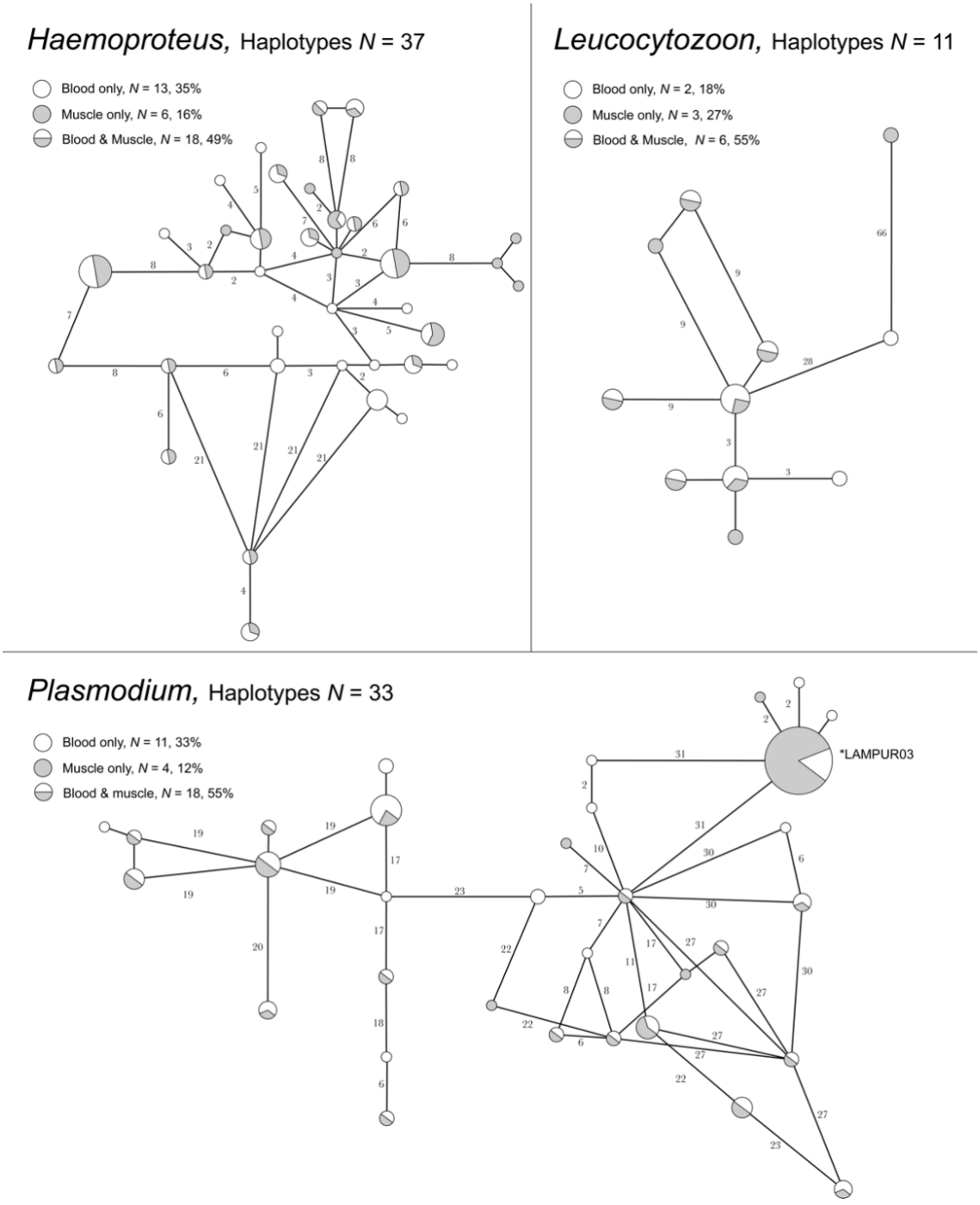

Figure 3. Minimum spanning networks for *Haemoproteus*, *Leucocytozoon* and *Plasmodium* of mtDNA cytb lineages recovered from blood and muscle source materials. Circle sizes represent frequency of the haplotypes. Mutational steps are noted at for all steps greater then 1. Reprinted with permission from (Harvey et al. 2017).

II.4 Discussion

Our study is the first to compare haemosporidian detection across source materials, using a broad sampling of avian species (Table 2). Although no differences in detection was found when I assessed all avian haemosporidian genera together, clear differences were found when examining each genus independently. Results recovered more detections, as well as higher genetic diversity, in *Haemoproteus* from blood source materials as compared to pectoral muscle. For *Plasmodium*, I found higher detections in pectoral muscle while recovering higher genetic diversity from blood source material. I found few *Leucocytozoon* detections across both blood and pectoral muscle starting materials. When taken together, results suggest that detection numbers and captured genetic diversity are not equivalent across pectoral muscle and blood starting material for each genus (Table 1). *Leucocytozoon* prevalence in our study is similar to previous work in nearby Cameroon and Gabon (Beadell et al. 2009).

Table 2. Host species taxonomic information including number of individuals sampled (N), locality sampled (C, P, D, L, T, A), frequency of detection across each genus *Haemoproteus* (H), *Leucocytozoon* (L), *Plasmodium* (P) and bioregion (B) sampled. Reprinted with permission from (Harvey et al. 2017).

Order	Family	Species	Common Name	Locality	N	H	L	P	B
Galliformes	Phasianidae	*Ptilopachus petrosus*	Stone Partridge	C (2)	2	0	0	2	S
Columbiformes	Columbidae	*Turtur abyssinicus*	Black-billed Wood-dove	C, P (2)	3	3		1	S
Charadriiformes	Glareolidae	*Rhinoptilus chalcopterus*	Bronze-winged Courser	C	1	0	0	0	S
	Jacanidae	*Actophilornis africanus*	African Jacana	A	1	0	0	0	G
Pelecaniformes	Ardeidae	*Butorides striata*	Striated Heron	C	1	0	1	0	S
Coraciiformes	Alcedinidae	*Alcedo quadribrachys*	Shining-blue Kingfisher	C	1	1	0	0	G
		Ceryle rudis	Pied Kingfisher	A	1	0	0	0	G/S
		Corythornis cristatus	Malachite Kingfisher	A, C	2	0	0	0	S
		Halcyon leucocephala	Gray-headed Kingfisher	C, P	2	1	0	0	S
		Halcyon malimbica	Blue-breasted Kingfisher	C, P	2	0	0	0	G
		Halcyon senegalensis	Woodland Kingfisher	T	1	1	0	0	G/S
		Ispidina picta	African Pygmy-kingfisher	C (5), P (4), D, T	11	1	0	0	S
	Meropidae	*Merops bulocki*	Red-throated Bee-eater	C (4), P	5	1	0	0	S
Piciformes	Indicatoridae	*Indicator indicator*	Greater Honeyguide	C	1	0	0	0	G/S
		Indicator minor	Lesser Honeyguide	C (2)	2	0	0	0	G/S
	Picidae	*Campethera punctuligera*	Fine-spotted Woodpecker	C (2)	2	0	0	0	S
	Ramphastidae	*Pogoniulus chrysoconus*	Yellow-fronted Tinkerbird	C	1	0	0	0	G
		Pogonornis dubius	Bearded Barbet	C	1	1	0	0	S
Passeriformes	Campephagidae	*Campephaga phoenicea*	Red-shouldered Cuckoo-shrike	P	1	0	0	0	S
	Cisticolidae	*Camaroptera brachyura*	Green-backed Camaroptera	C (3), P, T (2), D	7	0	0	5	S
		Camaroptera brevicaudata	Grey-backed Camaroptera	L, C, D	3	0	0	0	S
		Cisticola cantans	Singing Cisticola	C (2)	2	0	0	0	G
		Cisticola lateralis	Congo Whistling Cisticola	D	1	0	0	0	S
		Eremomela pusilla	Senegal Eremomela	C	1	1	0	0	G
		Prinia subflava	Tawny-flanked Prinia	C (4)	4	1	0	6	S
	Dicruridae	*Dicrurus adsimilis*	Fork-tailed Drongo	C	1	0	0	0	S
	Estrildidae	*Estrilda caerulescens*	Lavender Waxbill	C	1	2	0	0	G
		Estrilda melpoda	Orange-cheeked Waxbill	D	1	1	0	0	S
		Lagonosticta larvata	Black-faced Firefinch	C (3)	3	0	0	0	S
		Lagonosticta rubricata	African Firefinch	D (2)	2	0	0	3	S
		Lagonosticta senegala	Red-billed Firefinch	C, P (4)	5	0	0	1	S
		Pytalia melba	Green-winged Pytilia	P	1	1	0	1	G
		Spermestes cucullata	Bronze Mannikin	T, D	2	0	0	1	S
		Uraeginthus bengalus	Red-cheeked Cordonbleu	C (3)	3	2	0	1	S
	Leiothrichidae	*Turdoides reinwardtii*	Blackcap Babbler	C	1	2	0	0	G/S
	Malaconotidae	*Laniarius barbarus*	Common Gonolek	C (2), P	3	1	1	0	S
		Tchagra senegalus	Black-crowned Tchagra	P	1	0	0	2	S
	Monarchidae	*Terpsiphone rufiventer*	Black-headed Paradise-flycatcher	L	1	1	0	1	G
		Terpsiphone viridis	African Paradise-Flycatcher	C (3), P (2)	5	0	0	1	G/S
	Muscicapidae	*Cossypha albicapilla*	White-crowned Robin-chat	C (2)	2	0	0	0	G
		Cossypha niveicapilla	Snowy-crowned Robin-chat	C (6), T (2)	8	0	0	3	S
		Melaenornis edolioides	Northern Black-flycatcher	C	1	1	0	0	G
		Muscicapa aquatica	Swamp Flycatcher	C (2)	2	1	0	1	G/S
		Stiphornis erythrothorax	Western Forest Robin	L (5)	5	0	0	1	G
	Nectariniidae	*Chalcomitra senegalensis*	Scarlet-chested Sunbird	C (6), P, D	8	7	2	1	G
		Cinnyris cupreus	Copper Sunbird	D (3)	3	0	0	6	G

Table 2. Continued

Order	Family	Species	Common Name	Locality	N	H	L	P	B
Nectariniidae		*Cinnyris pulchellus*	Beautiful Sunbird	C (2), P (2)	4	0	0	5	S
		Cinnyris venustus	Variable Sunbird	C (2), T	3	0	1	3	G
		Cyanomitra olivacea	Eastern Olive Sunbird	L (5)	5	1	0	6	S
		Cyanomitra sp.		D	1	1	0	1	S
		Hedydipna collaris	Collared Sunbird	D (3)	3	0	0	5	S
		Hypergerus atriceps	Oriole Warbler	C	1	0	0	0	G
		Nectarina sp.		D	2	2	0	1	S
		Prinia erythroptera	Red-winged Prinia	D	1	0	0	1	G
	Passeridae	*Gymnoris dentata*	Bush Petronia	P	1	1	0	0	S
		Gymnoris pyrgita	Yellow-spotted Petronia	P	1	1	0	0	G
		Passer griseus	Northern Gray-headed Sparrow	P	1	1	0	1	G
	Pellorneidae	*Illadopsis puveli*	Puvel's Illadopsis	L	1	0	0	0	G/S
	Platysteiridae	*Batis senegalensis*	Senegal Batis	C	1	0	0	0	G/S
		Dyaphorophyia castanea	Chestnut Wattle-eye	L (2)	2	0	0	0	G/S
		Platysteira cyanea	Brown-throated Wattle-eye	C (5)	5	0	0	0	G/S
	Ploceidae	*Amblyospiza albifrons*	Grosbeak Weaver	T (2)	2	0	0	0	G
		Euplectes ardens	Red-collared Widowbird	D (2)	2	0	0	1	G
		Ploceus cucullatus	Village Weaver	T (2), D (4)	6	1	0	2	G
		Ploceus nigricollis	Black-necked Weaver	C (2), T (3)	5	1	0	2	G/S
		Ploceus ocularis	Spectacled Weaver	C, T (2)	3	0	0	3	S
	Pycnonotidae	*Andropadus virens*	Little Greenbul	L (2), D	3	0	0	0	G
		Bleda canicapillus	Gray-headed Bristlebill	L (5)	5	0	0	1	G
		Phyllastrephus albigularis	White-throated Greenbul	L (5)	5	0	0	1	G
		Phyllastrephus baumanni	Baumann's Greenbul	L	1	0	0	0	S
		Pycnonotus barbatus	Garden Bulbul	C (6), P, T, D	9	5	7	4	G/S
	Stenostiridae	*Elminia longicauda*	African Blue-flycatcher	C (2)	2	0	0	0	S
	Sylviidae	*Sylvietta brachyura*	Northern Crombec	D	1	0	0	0	S
		Sylvietta virens	Green Crombec	D (2)	2	2	0	1	S
	Turdidae	*Turdus pelios*	African Thrush	D	1	0	0	1	S
	Vangidae	*Prionops plumatus*	White Helmetshrike	C	1	0	1	0	S
	Viduidae	*Vidua sp.*		P, D	2	1	0	1	S

The localities previously sampled in Cameroon and Gabon included sites with tropical savanna climate (like Benin) but were predominantly tropical monsoon climates with higher overall precipitation with less variance in temperatures. *Leucocytozoon* vectors (Simuliidae: blackflies) are present in Benin (species inventories have been conducted for most countries with highest interest in medically relevant species), but the vectors have not been assessed for abundance or biogeographically across much of

Africa (Adler and Crosskey 2015). Blackfly presence is closely tied to rivers and other bodies of water and therefore is spatially restricted (Sutcliffe 1986). Species richness has been associated largely with temperature and stream discharge, and it has been suggested that they are rarer in tropical climates (Young et al. 1993; Ya'cob et al. 2016). Therefore, I attribute the *Leucocytozoon* prevalence in our study to the timing of sampling (early in the rainy season) and the relative climatic characteristics of the sampling regions and associated vector abundance. The differences in life stage development and transmission seasonality vary as I assess each genus, and within these differentiations, I believe lay three possible explanations for the patterns in detections and haplotype diversity across genera and source material. First, the seasonality of transmission period and the associated length of these infection periods vary across parasite host genera and even within closely related parasites lineages, indicating that parasite lineages may exhibit evolutionary adaptions to transmission strategy (Valkiūnas et al. 2004; Pérez-Rodríguez et al. 2015). Avian hosts spend more time in chronic or latent infection periods as compared to prepatent or acute infection periods, hence leading to high detections in pectoral muscle (Valkiūnas 2005). However, acute parasitemia is generally under-sampled due to reduced host activity (50% less activity e.g., highly reduced flight, foraging, and breeding activity, as compared to uninfected birds) (Valkiūnas 2005; Mukhin et al. 2016). Birds that can activate a sufficient immune response move from acute to the chronic phase (low parasitemia) and then to latent infections; otherwise, mortality from acute infection likely occurs (Valkiūnas 2005). Infected birds with low-level chronic infections are the most detectable via mist netting given the relatively

benign symptoms, which do not hinder normal physiological maintenance movements of birds (i.e., foraging, nesting, breeding). These low-level infections are detectable in both blood and fixed tissue and thereby resulted in high detection rates in both blood and muscle. Additionally, the chronic stage is longer than other stages (months for *Haemoproteus* and *Leucocytozoon*; up to a year for *Plasmodium*), making this stage the most detectable in wild populations (Valkiūnas 2005). Moreover, older birds have been shown to have higher parasite prevalence than younger birds, possibly due to increased hormone levels and the associated incidence of relapse combined with the exposure to new seasonal infections (Greiner and Mundy 1979; Deviche et al. 2001).

Second, life stages may impact haemosporidian detection probability across source starting material. The location (i.e., in blood, pectoral muscle, or other organs) of the parasite across life stages is known from experimental infection studies (Fallis and Bennett 1960; Hepler et al. 1966; Khan and Fallis 1970; Atkinson et al. 1986; Atkinson et al. 1988; Zehtindjiev et al. 2008; Valkiūnas et al. 2015), although these stages are described from a small proportion of species from each haemosporidian genus (Valkiūnas 2005; Bensch et al. 2009). So, while informative, there is a lack of broad knowledge regarding how consistent life stage locations (and durations in each) are across the high diversity of lineages recovered, particularly if the transmission strategies show plasticity as previously seen in discrete lineage clades (Pérez-Rodríguez et al. 2015). Regardless, I expect overall lower detections in blood for birds unless specifically targeting the timing of active transmission periods (i.e., primary parasitemia: acute and chronic), when meronts and gametocytes are found in the circulating blood. Our high

detection rates in pectoral muscle may be due to parasite recovery from the tissue stages of infections (i.e., located in the endothelial lining of capillary cells or skeletal muscle) which may be confounded by possible carryover of circulating blood in muscle (Valkiūnas 2005). Further, some of the muscle tissue only detections may have captured abortive infections (i.e., dead end infections) (Markus 2011). Abortive development has been found to result in morbidity and acute disease as a result of the damage from meronts to internal organs. Consequently, this has most commonly been detected in exotic captive species, being difficult to capture and assess in wild populations (Donovan et al. 2008; Olias et al. 2011; Cannell et al. 2013). Diagnosing abortive development requires necropsy and histology examinations, though molecular sequences can be examined for signals of abortive development by looking at host associations along with haemosporidian genera host specificity. I only found one likely detection of abortive development, where a Striated Heron (*Butorides striata*) recovered a lineage most homologous to a *Haemoproteus* lineage previously detected in the Common Blackbird (*Turdus merula*).

Third, resident birds (i.e., non-migratory or narrowly intra-continental migrants) in tropical climates engage in diverse breeding strategies. Timing of tropical breeding is mostly related to precipitation patterns (i.e., rainy season occurrence and the resulting humid period) and subtle photoperiodic cues, though some avian species have continuous or opportunistic breeding strategies which are often related to non-seasonal rainfall (Hau et al. 2008). Relapses of infection have been associated with gonadal, stress, and pineal hormones which are activated by photoperiod and associated breeding

cues, though no conclusive mechanism has been discovered (Desser et al. 1968; Applegate and Beaudoin 1970; Valkiūnas et al. 2004; Cornelius et al. 2014). Additionally, if temperature and relative humidity are appropriate for parasite development and vectors are available, transmission may be occurring year-round (Hasselquist et al. 2007; Sorensen et al. 2016). However, extreme dry periods characterizing the tropical savanna climate of our Benin sampling localities make year-round transmission unlikely to occur. Furthermore, parasite persistence is sustained via these relapses, which then enables the infection of nestlings and juveniles (Bennett and Cameron 1974; Greiner and Mundy 1979; Deviche et al. 2001; Dunn et al. 2016). Yet, sampling for this study took place before the rainy season in the northern sampling localities, so presumably breeding for most birds was not yet initiated and consequently chronic, latent, and relapse infections are most likely detectable during this period. Sampling for southern localities took place in the middle of the first of two annual rainy seasons suggesting some birds had commenced breeding for the season and that relapses as well as new primary parasitemia periods are likely to be detected in these birds.

Previous source material comparison studies found no significant differences for haemosporidian parasite detection (Ramey et al. 2013; Svennson-coehlo et al. 2016). These investigations were limited to single host species studies, recovery of a single primary haemosporidian genus (*Leucocytozoon* and *Plasmodium*, respectively), and the concordant host-parasite relationship bias. Given that not all avian hosts are equally susceptible to *Haemoproteus*, *Leucocytozoon*, and *Plasmodium*, broader comparisons and extrapolations can therefore not be made without including greater host taxon

sampling. Avian host susceptibility also varies across each parasite genus, with *Haemoproteus* and *Leucocytozoon* species having higher host specificity (phylogenetically host family restricted) as compared to *Plasmodium* which exhibits much lower host specificity (Bennett et al. 1982; Atkinson and Van Riper 1991; Beadell et al. 2009; Lutz et al. 2015). Host susceptibility is also affected by different vector associations, as each parasite genus is vectored by different dipteran groups whose presence and abundance have differing ecological constraints and timing of emergence (Atkinson and van Riper 1991; Valkiūnas 2005). Therefore, our results indicate that each avian haemosporidian genus should be assessed independently. Blood starting source material was the most informative for *Haemoproteus*, providing a higher number of detections and greater genetic diversity. Both blood and pectoral muscle were informative for *Plasmodium*, with pectoral muscle providing more detections and blood providing higher genetic diversity. When examining just one source material, there may be a potential loss of genetic diversity and possible prevalence underestimation, particularly when pectoral muscle is the single source material. I realize that many studies may not have the opportunity to collect tissues other than blood, but I encourage those associated with museums to employ broader collections of source materials when performing broad diversity studies of hosts (as was the impetus for our sampling in Benin).

We conclude by suggesting that when investigations of avian haemosporidians are initiated, study design needs to take into consideration three factors. First, the genus (genera) of interest should inform the choice of sampling material. Second, the timing of

sampling has two important facets, as seasonality of transmission is dependent on availability and abundance of vectors, and the periods shortly after host breeding season or rainy season are likely best for capturing the chronic infection stage. Clearly, more vector assessment is needed for all haemosporidian vectors across Africa (Dipterans: Ceratopogonidae, Culicidae, Hippoboscidae, and Simuliidae) in order to disentangle parasite transmission strategies (i.e., changes in length of chronic periods) from limitations of vector seasonality. Third, the climate of the sampling location (i.e., tropical versus temperate) as climate is informative for seasonality of hosts and vectors. The heterogeneity in detection I found across source materials at the genus level indicates the importance of source material selection for parasite studies of not only parasite diversity and ecology, but for deeper understanding of ecological patterns using comparative studies.

CHAPTER III

HOST ASSOCIATIONS AND CLIMATE INFLUENCE AVIAN HAEMOSPORIDIAN DISTRIBUTIONS IN BENIN

III.1 Introduction

Species are bound geographically by various factors such as oceans, mountains, or islands; however, vector borne parasites have additional constraints, which include distributions of their arthropod vectors, competent host availability, and other host biotic factors (e.g., age, density, sex, life history characteristics). Furthermore, arthropods that vector parasites are restricted by their own environmental constraints including temperature, precipitation, and humidity, all of which combine to determine life cycle events (Gage et al., 2008). As such, the prevalence of vector borne parasites is dependent on both vector abundance and the ecological conditions required by their vectors (van Riper III et al., 1986). Integrating bioclimatic factors as a proxy for vector distributions seems a critical consideration for studies seeking to determine the strength of association with parasites across hosts and geography.

Given the predicted effects of global climate change and associated changes in diurnal fluctuations I expect to see the following effects on vectors and ultimately on parasite distributions 1) expanded vector distributional ranges and associated altered parasite distributions, 2) extension of the seasonal activity of both vectors and parasites thus extending transmission periods, and 3) increased introductions where vector/parasite fauna are novel (González et al., 2010; Garamszegi, 2011; Caminade et

al., 2014). Range extensions and shifts will increase the introduction of parasites, and thus increase the infection of novel host species (Garamszegi, 2011; Altizer et al., 2013; Loiseau et al., 2013); the decimation of the Hawaiian avifauna remains the classic example of this effect (van Riper III et al., 1986). Increases in temperature will also directly impact parasite development, which is temperature limited (Valkiunas, 2005; LaPointe et al., 2010). For example, studies on *Plasmodium* indicate that increases in temperature will increase the rate of development and thus, result in increases in prevalence and their impact on host populations (LaPointe et al., 2010). Therefore, determining current distributions and constraints of taxa is necessary to better predict future changes in distributions.

Haemosporidians are protozoan blood parasites infecting vertebrate taxa including reptiles, mammals, and birds, with transmission occurring via blood sucking dipterans. Avian haemosporidians range nearly worldwide and consist of the genera *Haemoproteus* (including *Parahaemoproteus*), *Leucocytozoon*, and *Plasmodium*. Each of these genera is vectored by a different suite of dipterans and these dipterans have varying ecological constraints to include moisture (ranging from humidity, moisture content of soil, precipitation, pooling of water, and running streams) and temperature regimes, which increases the complexity for understanding and modeling haemosporidian distributional limits. It is known that *Haemoproteus* is vectored by biting midges (Ceratopogonidae). Biting midges exhibit highly varied habitat preferences with most biting midges inhabiting damp sand or soil, particularly along rivers and marshes, while they can also be found breeding in detritus, pools, and streams

(Meillon and Wirth, 1991; Mellor et al., 2000; Meiswinkel et al., 2004). The primary vectors of *Haemoproteus*, Louse flies (Hippoboscidae), are highly restricted to host and their larvae are laid in or near bird nests; given that louse flies do not lay eggs and instead deposit larvae directly, their moisture constraints are limited (Lamerton, 1965). *Plasmodium* is vectored by mosquitos (Culicidae) which vary in breeding site preference across species and have high moisture requirements for breeding and emergence, yet are preferentially found in swamp and upland forest versus open habitats (Njabo et al., 2009). *Leucocytozoon* is vectored by black flies (Simuliidae) and requires some degree of flowing water for egg laying and is thereby primarily found near rivers and other bodies of water (Sutcliffe, 1986).

Although there is basic knowledge of which insect vectors transmit haemosporidian genera, the specific vector species for most haemosporidian species remains unknown, and overall there are few studies linking the vector to both the vertebrate host and the haemosporidian (Malmqvist et al., 2004; Martínez-De La Puente et al., 2011; Santiago-Alarcon et al., 2012). Additionally, there appears to be a great deal of undiscovered Haemosporidia diversity, as new clades and even genera are being described with regularity (Bertram et al., 2017). Further, many geographic regions and host taxa have not been sampled for haemosporidians, thereby impacting our ability to map parasite associations and distributions, which in turn impacts our ability to predict range shifts under climate change scenarios. So, although avian haemosporidians have become a model parasite system, much remains to be learned before it can reliably be used to track climate change impacts on communities.

Several studies have attempted to assess the importance of environmental variables in predicting the distribution and prevalence of haemosporidian parasites, although many of these studies have been on single or a limited number of host species and their associated haemosporidians (Loiseau et al., 2012b, 2013). Researchers examining the distribution of haemosporidians across avian communities along two mountain ranges found that indeed each parasite genus demonstrated some environmental preferences, including preference across elevational gradients (Illera et al., 2017). While examining a single widespread tropical forest species, researchers found that maximum temperature of the warmest month was the strongest predictor of *Plasmodium* prevalence (Sehgal et al., 2011). These studies indicate the need for assessment within and across avian/haemosporidian communities.

Climate change is already having an effect: mortality due to chronic malaria was recently documented in a Common Loon (*Gavia immer*) individual in New Hampshire, USA. This is a previously monitored and believed to be uninfected Nearctic breeding species, which indicates a possible shift in the *Plasmodium* range or a shift in host species susceptibility, as suggested by the high level of virulence detected histologically (Martinsen et al., 2017). Indeed, the northern latitudinal limit of *Plasmodium* has expanded and is associated with temperature changes; further latitudinal expansions poleward are predicted, as are elevational expansions in montane areas (Loiseau et al., 2012a, 2013). Much of haemosporidian diversity remaining not sampled along with a lack of understanding of vector associations; these limitations present challenges in fully

characterizing the dimensionality of the relationship of parasites and the associations with host and bioclimatic data variables.

Here we examine haemosporidian distributions in the western African country of Benin, previously a subset of this data was assessed for detections across source materials (Harvey and Voelker, 2017). Our sampling is ideal due to the small geographic scale, yet ecologically diverse sites. Benin is located within the Sudanian bioregion and across two contrasting ecoregions: west Sudanian savanna and Guinean forest-savanna (Linder et al., 2012; Dinerstein et al., 2017). The localities sampled have generally tropical savanna climates (Peel et al., 2007), with mean temperatures above 18^{o}C and a marked dry season, but they are differentiated by length and timing of wet seasons which correspond to Northern (arid) and Southern (generally moist) localities (Figure 4). Both of the west Sudanian savanna ecoregion localities, Chutes de Koudou and Point Triplo, are characterized by a long dry season with a pronounced single wet season, which occurs from June to September (http://worldclim.org; Hijmans et al., 2005). All of the Guinean forest-savanna localities, Dogo Forest, Lama Forest, Lake Toho, and Abomey Calavi, have two wet seasons. Our sampling (Figure 4) thereby provides a latitudinal gradient framework from the northern west Sudanian savanna localities to the southern Guinean forest savanna localities. Here, we examine avian haemosporidians, and assess how their host relationships and environments are shaping the resulting diversity and distributions in Benin. We seek to elucidate: 1) host-parasite associations across a climatic gradient, 2) the association of haemosporidian distributional patterns with bioclimatic variables, and 3) how the associations of 1 and 2 inform the observed

distributional patterns. We associate these objectives across scales, addressing the bioregion, the ecoregions, and the localities sampled.

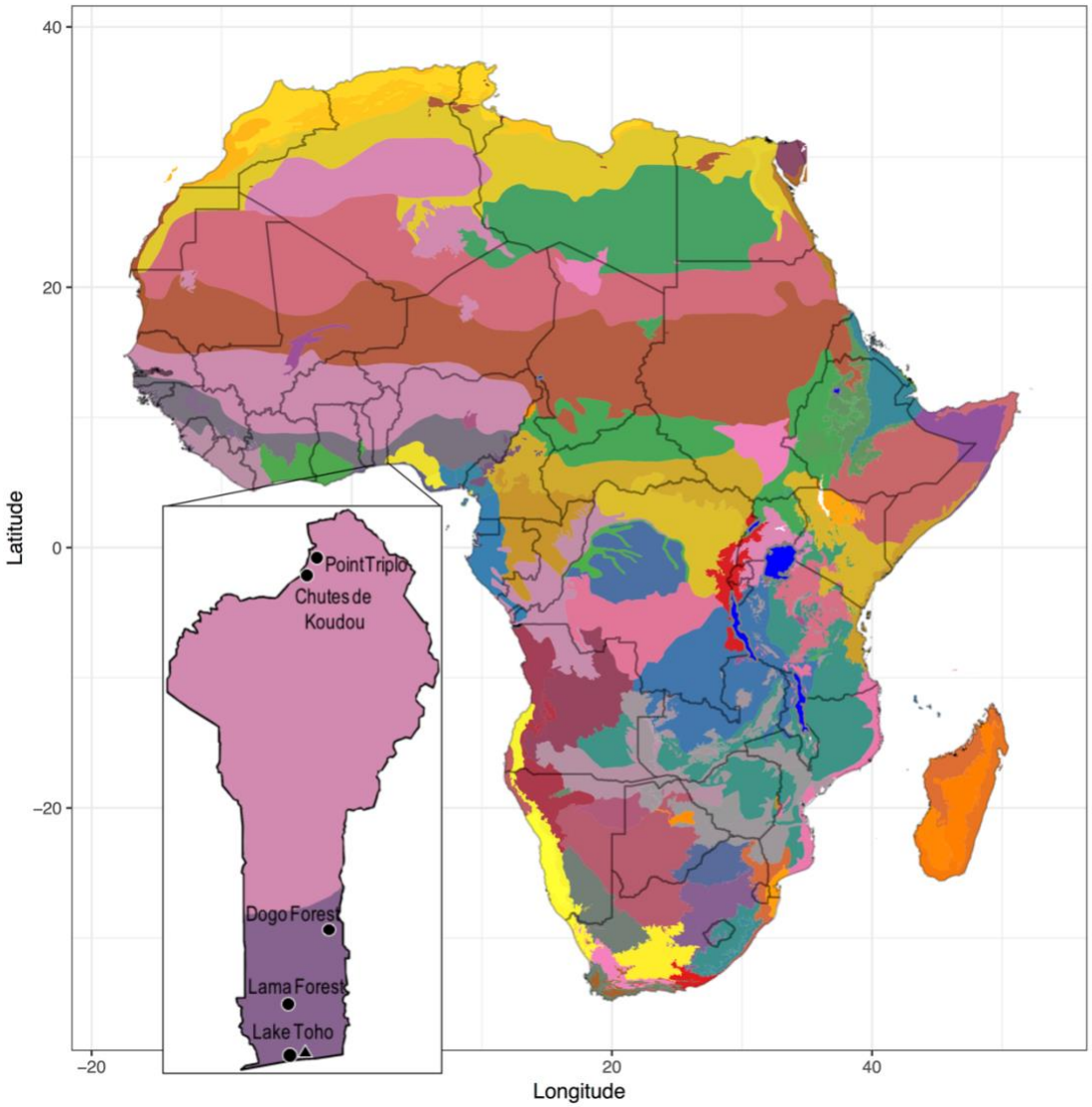

Figure 4. Map demonstrating diversity of ecoregions across African continent (n = 110). Inset map of Benin's ecoregions (n = 2), West Sudanian savanna in lilac (north), and Guinean forest savanna in medium purple (south). Sampling localities (black circles) and merged locality denoted (triangle).

III.2 Methods

III.2.1. Sampling localities and avian sampling

Birds were sampled between May and June of 2010, as described in Harvey and Voelker (Harvey and Voelker, 2017). Avian sampling includes the 199 individual birds previously sampled (from blood and pectoral muscle source materials) and the addition of 23 individuals for which we have blood (N = 13) or pectoral muscle (N = 10) sampled. From here on all 222 individual birds are addressed for haemosporidian detection, including all detections recovered from both source materials.

Briefly, birds were sampled at six localities across Benin. Sampling at the northern localities, Point Triplo and Chutes de Koudou localities, occurred between May 21st – May 28th. All southern localities (Dogo Forest, Lama Forest, Lake Toho and Abomey-Calavi) were sampled from June 1st – June 10th (Figure 4). Only three samples were collected in Abomey-Calavi, which lies 20 km east of Lake Toho, and given the similar habitat and bioclimatic variables these samples were grouped together.

All voucher specimens collected are accessioned in the Biodiversity and Research Teaching Collections, at Texas A&M University. We collected all specimens under protocols approved by the Institutional Animal Care and Use Committee at Texas A&M University.

III.2.2 Molecular assessment of avian haemosporidians

We followed molecular protocols previously described in Harvey and Voelker (Harvey and Voelker, 2017). We targeted a fragment of the mitochondrial (mtDNA) Cyt *b* gene using multiple primer pairs previously published in Drovetski et al. (2014). This

fragment encompassed the entire 479 base pairs of the standard gene region collected in the MalAvi avian haemosporidian database (Bensch et al., 2009). Collectively, the primers from Drovetski et al. (2014) amplify all three genera of avian haemosporidians: *Haemoproteus* (to include subgenera *Parahaemoproteus* and *Haemoproteus*), *Leucocytozoon*, and *Plasmodium.*

Sequences were verified for quality base by base and aligned by eye using Geneious 6.1.8 (http://www.geneious.com, Kearse et al., 2012). Multiple infections were determined by the presence of multiple peaks on both chromatograms at one or more base positions (Harvey and Voelker, 2017). Due to the high probability that sequences with less than three DNA positions displaying multiple peaks were sequencing errors and not true multiple infections (Szymanski and Lovette, 2005), we treated these as single infections. After being verified with criteria for peak similarity, confidence score and visual assessment we processed double infection data (n=5) with assigned IUPAC ambiguity codes. We then reconstructed single infection haplotypes (Browning and Browning, 2011) using Phase 2.1 (Stephens et al., 2001) as implemented in DnaSP 25.10.1 (Librado and Rozas, 2009) along with all similar (≥ 97% Blast match) of Benin and MalAvi data.

All data were identified to genus by use of the MalAvi blast (Version 2.2.8, Bensch et al., 2009) and NCBI BLAST (Altschul et al., 1990) functions. Criteria of one base pair was followed to differentiate genetic lineages (i.e. unique haplotype) (Bensch et al., 2000, 2004). Sequences were assigned the MalAvi lineage name (if identified as a

complete match) or identified as novel detections (Genbank accession numbers "pending").

III.2.3 Delimitation of parasite clades of interest

To assess diversity of haemosporidians we need to determine to biological units we are addressing. The delineations for haemosporidian evolutionary units, species and lineages, are poorly understood. Lineage formations have been found to develop primarily through host switching and allopatric speciation (Ricklefs and Fallon, 2002; Ricklefs et al., 2014). Species descriptions have been primarily based on morphological descriptions of circulating red blood cell life stages (meronts and gametocytes) using microscopy of blood smears, while some species description are across all life stages (this is methodologically more challenging and therefore rare) (Valkiunas, 2005). Currently there are 220 morphologically described haemosporidian species (MalAvi 2.3.3). Molecular determinations, using the single nucleotide substitutions within the 479 base pair Cyt b region, have yielded 2,876 molecular lineages thus far (MalAvi 2.3.3). The one base pair delimitation for genetic lineages has been accepted as the best practice (Bensch et al., 2000, 2004) and is supported by the fact that Haemosporidian lineages have not shown evidence of recombination events (Joy, 2003; Bensch et al., 2004, 2009).

Delimitations of species are more complex and not standardized across haemosporidian taxa. The sequence divergence between the human *Plasmodium falcip*arum and the chimpanzee *P. reichenowi* is 3.3% sequence divergence within human parasite *Plasmodium falciparum* parasites (N = 96) reaches 0.2% across all of

Cyt b, consisting of six haplotypes with single base pair substitutions (Joy, 2003). When examining the same *P. falciparum* data set restricted down to the 505 base pairs assessed in this study pairwise sequence divergence reaches 0.4% consisting of four haplotypes with single base pair substitutions (Joy, 2003). Avian *P. relictum* from the Hawaiian Islands was previously believed to show no sequence divergence while globally distributed *P. relictum* lineages demonstrate a 7.6% sequence divergence (Beadell et al., 2006). A study by Jarvi et al. (Jarvi *et al.* 2013) demonstrated 23 variant haplotypes recovered from Hawaiian *P. relictum* using deep sequencing. However, these variants had low coverage (ranging from one to nine) and an average read depth of 4.6x. A number of these variants included non-synonymous substitutions or resulted in stop codons, suggesting that these variants may be a result of sequencing error or random mutations which are biologically not meaningful; thus, supporting the genetic conservatism previously detected in Hawaiian *P. relictum*. This variability in the divergence rates across groups supports the argument to range criteria (1 -5 %) for species delimitations and use other characteristics as support, including morphology where available, host association, and locality (Outlaw and Ricklefs, 2014).

Here we are not attempting to determine species as morphological data is missing. Instead we address phylogenetic clades of haemosporidians to determine patterns of association across hosts and climate. Clades are selected with the criteria of being reciprocally monophyletic and having a within group pair wise sequence divergence of less than 5.5% though most are more conservative. Pairwise sequence divergence was measured using MEGA version 7.0.14 (Kumar et al., 2016).

III.2.4 Bioclimatic Data

We used the WorldClim 2 database and the ENVIREM data set and their bioclimatic and topographic variables as these have been determined to be biologically important for species distributions (Fick and Hijmans, 2017; Title and Bemmels, 2017). WorldClim 2 data are based on interpolated minimum, maximum, and monthly averages of precipitation and temperature, and were collected from 1970-2000 at ~1km^2 resolution. We included monthly variables such as precipitation, water vapor pressure, wind speed, and temperature maximum and minimum for the second quarter (April, May and June) as this is reflective of vector emergence, infection, and resulting host infection for the collecting period in which we sampled (Table 3). From the EVIREM data set, we included additional variables associated with measures of potential evapotranspiration, climate moisture index, and an aridity index, and additional climate and topographic data. ENVIREM data was collected from ~1960-1990 and also at ~1 km^2 resolution. Our initial data climate set contains 53 variables (Table 3).

Table 3. Bioclimatic variables included used in the analysis. Variables retained and used in the heat map analysis notated with an *. Variable name, explanation of variable if needed, units, and source (WorldClim2 or ENVIREM).

Variable Name Used	Explanation if needed	Units	Source
Precipitation June*		mm	WorldClim V2
Climatic Moisture Index*	Metric of relative wetness and aridity		ENVIREM
Precipitation May*		mm	WorldClim V2
EmbergerQ*	Metric designed to differentiate among Mediterranean type climates		ENVIREM
Precipitation Driest Month*		mm	WorldClim V2
Precipitation Driest Q*		mm	WorldClim V2
Isothermality*	Mean diurnal range/Temperature annual range	Mean Diurnal Range/Temperature Annual Range	WorldClim V2
Annual Precipitation*		mm	WorldClim V3
Precipitation Warmest Q*		mm	WorldClim V2
Precipitation April*		mm	WorldClim V2
Precipitation Coldest Q*		mm	WorldClim V2
PET Wettest Q*	Mean monthly PET of wettest quarter	mean monthly PET of wettest quarter (mm/month)	ENVIREM
PET Driest Q*	Mean monthly PET of driest quarter	mean monthly PET of driest quarter (mm/month)	ENVIREM
Annual PET*	Annual potential evapotranspiration	mm/year	ENVIREM
PET Warmest Q*	Mean monthly PET of warmest quarter	mean monthly PET of warmest quarter (mm/month)	ENVIREM
Temp Annual Range*	Max Temp Warmest month- Min Temp Coldest month	℃	WorldClim V2
PET Coldest Q*	Mean monthly PET of coldest quarter	mm/month	ENVIREM
Mean Diurnal Range*	Mean of monthly (max temp - min temp)	℃	WorldClim V2
PET Seasonality*	Monthly variability in potential evapostranspiration (	mm/month	ENVIREM
Mean Temp Wettest Q*		℃	WorldClim V2
Precipitation Wettest Mo.*		mm	WorldClim V2
Temp Avg June*		℃	WorldClim V2
Temp Avg April*		℃	WorldClim V2
Temp Avg May*		℃	WorldClim V2
Aridity I Thornwaite*	Index of the degree of water deficity below water need	-	ENVIREM
Temp Seasonality*	Standard deviation of temperature	℃	WorldClim V2
Continentality*	Average temp. of warmest month - average temp. of coldest month	℃	WorldClim V2
Precipitation Wettest Q*		mm	WorldClim V2
Min Temp Warmest Mo.*		℃	ENVIREM
Temp Min June*		℃	WorldClim V2
Wind April*		m s-1	WorldClim V2
Solar Radiation*		kJ m-2 day -1	WorldClim V2

Table 3. Continued

Variable Name Used	Explanation if needed	Units	Source
Precip Seasonality*	Coefficient of variation	-	WorldClim V2
Wind May*		m s-1	WorldClim V2
Temp Min May*		℃	WorldClim V2
Temp Min April*		℃	WorldClim V2
Wind June*		m s-1	WorldClim V2
Annual Mean Temp		℃	WorldClim V2
Max Temp Warmest Mo		℃	ENVIREM
Min Temp Coldest Mo		℃	WorldClim V2
Mean Temp Driest Q		℃	WorldClim V2
Mean Temp Warmest Q		℃	WorldClim V2
Mean Temp Coldest Q		℃	WorldClim V2
Growing Deg Day 0	Sum of mean monthly temperature for months with mean temperature greater than 0℃ multiplied by number of days		ENVIREM
Growing Deg Days 5	Sum of mean monthly temperature for months with mean temperature greater than 5℃ multiplied by number of days		ENVIREM
Max Temp Coldest		℃	ENVIREM
Thermicity Index	Sum of mean annual temp., min. temp. of coldest month, max temp. of coldest month X10	℃	ENVIREM
Temp Max April		℃	WorldClim V2
Temp Max May		℃	WorldClim V2
Temp Max June		℃	WorldClim V2
Vapor Apr		kPa	WorldClim V2
Vapor May		kPa	WorldClim V2
Vapor June		kPa	WorldClim V2

We used partial least squares regression (PLSR) to model the relationships of bioclimatic variables (as predictor variables) to haemosporidian richness and prevalence recovered across localities (response variables). PLSR multivariate regression is best suited for a high number of predictor variables and a small number sampled response variables, responding favorably to highly correlated data (Mevik and Wehrens, 2007). The high number of predictor variables (n = 53; Table 3) are transferred from linear correlations of the factors to latent variables (or components), for which the covariance between response and predictor variables is maximized. We conducted the PLSR as implemented in the *pls* package (Mevik and Wehrens, 2007) in the R software version 3.2.2 (R Core Team, 2012) and examined our initial bioclimatic data set of 53 predictor variables along with measures of richness and prevalence for each haemosporidian genus across sampling localities as response variables. The SIMPLS algorithm was used in PLS regression given the high number of predictor variables (de Jong, 1993). The resulting number of latent variables were reduced using a leave one out cross-validation method. Prevalence was calculated as the proportion of individuals infected for each genus across each sampling locality. Richness was calculated as the number of unique lineages recovered across each sampling locality. Bioclimatic data were log(x+1) transformed. Response variables were square root transformed for counts and logit transformed for proportions. We used the resulting root mean squared error of prediction (RMSEP) validation results to determine the number of latent variables as most predictive for each response variable (Figure 5). We then used the regression coefficients for each response variable across latent variables and determined the most

significant predictors for each association. All bioclimatic variables (predictors) with significant values were selected for the number of components with the corresponding lowest RMSEP value across all response variables, resulting in n=37 predictor variables (Table 4). The significance of the predictors was determined by taking the squared value of the coefficient and determining those greater then 1/k (where k is the number of predictor variables).

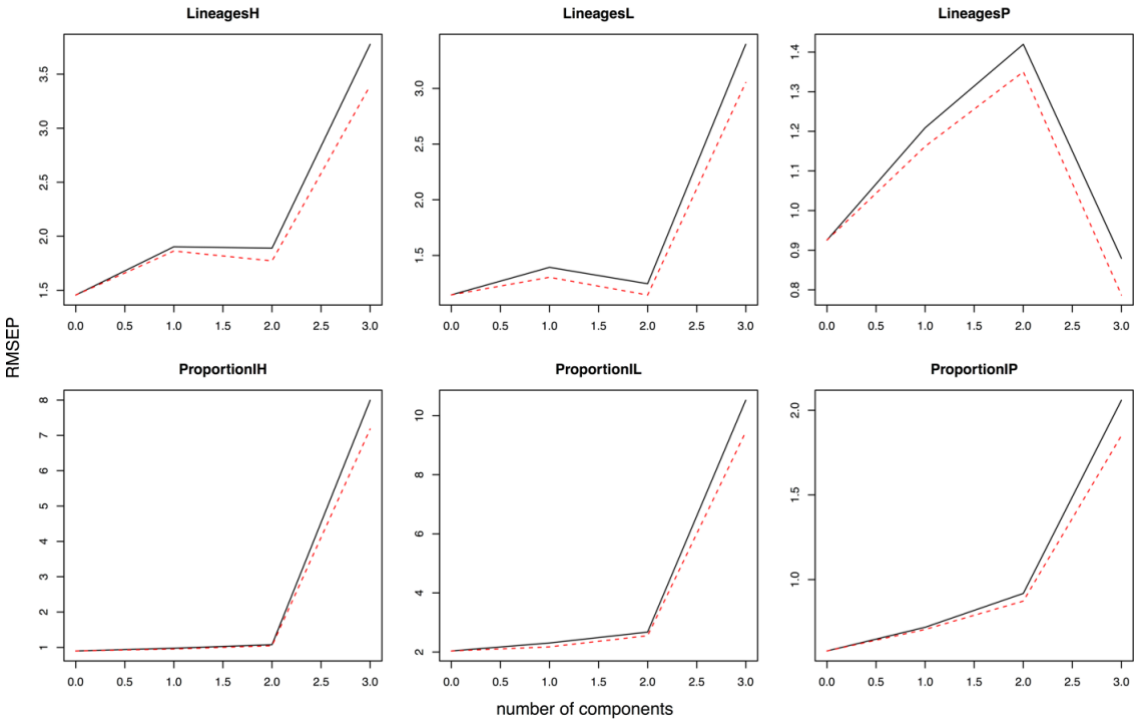

Figure 5. Root mean squared error of prediction (RMSEP) validation results to determine the number of latent variables as most predictive for each response variable (Number of lineages for each parasite genus and proportion infected for each genus).

Table 4. Regression coefficients for each response variable across latent variables which were determined significant predictors using the root mean squared error of prediction values.

	Haemoproteus				***Leucocytzoon***				***Plasmodium***				
	Richness		Prevalence		Richness			Prevalence	Richness			Prevalence	
	LV1	LV2	LV1	LV2	LV1	LV2	LV3	LV1	LV2	LV1	LV3	LV1	LV2
	47.28	29.82	40.42	9.75	49.09	32.5	40.83	19.97	61.09	11.49	41	61.12	7.51
Precipitation June	0.004	0.006	0.001	0.001	0.003	0.004	**5.626**	0.003	0.001	-	0.006	0.001	0.001
Climatic Moisture Index	0.005	0.009	0.002	0.001	0.003	0.006	**0.960**	0.004	0.002	-	-	0.001	0.001
Precipitation May	0.013	**0.026**	0.004	0.002	0.008	0.017	**0.441**	0.011	0.006	0.001	0.002	0.003	0.002
EmbergerQ	0.017	**0.035**	0.005	0.003	0.011	**0.023**	**2.218**	0.014	0.009	0.002	-	0.003	0.002
Precipitation Driest Mo	**0.112**	**0.307**	**0.036**	0.013	**0.072**	**0.202**	0.003	**0.092**	**0.093**	0.011	**0.081**	**0.023**	0.011
Precipitation Driest Q	**0.269**	**0.758**	**0.088**	**0.029**	**0.173**	**0.498**	**2.775**	**0.222**	**0.233**	**0.026**	**0.275**	**0.055**	**0.027**
Isothermality	-	-	-	-	-	-	**0.132**	-	-	-	0.001	-	-
Annual Precipitation	0.001	-	-	-	0.001	-	**0.897**	0.001	-	-	0.002	-	-
Precipitation Warmest Q	0.016	**0.040**	0.005	0.002	0.011	**0.026**	**0.536**	0.014	0.011	0.002	**0.017**	0.003	0.002
Precipitation April	**0.069**	**0.164**	**0.022**	0.010	**0.044**	**0.107**	**0.914**	**0.057**	**0.045**	0.007	0.057	0.014	0.008
Precipitation Coldest Q	0.011	**0.497**	0.004	0.119	0.007	**0.338**	**0.356**	0.009	**0.494**	0.001	**0.496**	0.002	**0.043**
PET Wettest Q	-	-	-	-	-	-	**1.203**	-	-	-	**0.005**	-	-
PET Driest Q	-	0.001	-	-	-	0.001	**0.651**	-	-	-	0.001	-	-
Annual PET	0.001	0.002	-	-	0.001	0.002	**0.390**	0.001	0.001	-	-	-	-
PET Warmest Q	0.001	0.003	-	-	0.001	0.002	**0.485**	0.001	0.001	-	-	-	-
Temp Annual Range	0.005	0.014	0.002	0.001	0.003	0.009	**0.172**	0.004	0.004	0.001	0.002	0.001	0.001
PET Coldest Q	0.002	0.005	0.001	-	0.001	0.003	**0.300**	0.001	0.001	-	-	-	-
Mean Diurnal Range	0.004	0.011	0.001	0.001	0.003	0.007	**0.523**	0.003	0.003	-	-	0.001	-
PET Seasonality	0.006	0.018	0.002	0.001	0.004	0.012	**0.609**	0.005	0.006	0.001	0.001	0.001	0.001

Table 4. Continued

	Haemoproteus				***Leucocytzoon***				***Plasmodium***				
	Richness		Prevalence		Richness		Prevalence		Richness			Prevalence	
	LV1	LV2	LV1	LV2	LV1	LV2	LV3	LV1	LV2	LV1	LV3	LV1	LV2
	47.28	29.82	40.42	9.75	49.09	32.5	40.83	19.97	61.09	11.49	41	61.12	7.51
Temp Avg May	-	0.001	-	-	-	0.001	**0.024**	-	-	-	-	-	-
Aridity I Thornthwaite	0.005	0.018	0.002	-	0.003	0.012	**1.163**	0.004	0.006	0.001	0.015	0.001	-
Temp Seasonality	0.008	**0.023**	0.003	0.001	0.005	0.015	**0.024**	0.007	0.007	0.001	0.007	0.002	0.001
Continentality	0.006	0.013	0.002	0.001	0.004	0.009	**0.225**	0.005	0.003	0.001	0.001	0.001	0.001
Precipitation Wettest Q	0.002	0.012	0.001	-	0.001	0.008	**4.938**	0.002	0.005	-	**0.029**	-	-
Min Temp Warmest Mo	-	-	-	-	-	-	**0.027**	-	-	-	-	-	-
Temp Min une	-	-	-	-	-	-	**0.121**	-	-	-	0.001	-	-
Wind April	0.001	0.006	-	-	0.001	0.004	**1.218**	0.001	0.002	-	0.009	-	-
Solar Radiation	-	-	-	-	-	0.000	**0.052**	-	-	-	-	-	-
Precipitation Seasonality	0.009	**0.030**	0.003	0.001	0.006	0.020	**3.380**	0.008	0.010	0.001	**0.031**	0.002	0.001
Wind May	0.003	0.008	0.001	-	0.002	0.006	**1.789**	0.002	0.003	-	**0.012**	0.001	-
Temp Min May	-	0.001	-	-	-	-	**0.110**	-	-	-	0.001	-	-
Temp Min April	-	0.001	-	-	-	0.001	**0.115**	-	-	-	0.001	-	-
Wind June	0.005	0.017	0.002	-	0.003	0.011	**3.713**	0.004	0.006	0.001	**0.025**	0.001	-

III.2.5 Phylogenetic analysis

We created the avian host phylogeny using birdtree.org, which uses the comprehensive avian phylogeny recovered by Jetz et al (2012) along with a fossil calibrated backbone phylogeny Hackett et al. (2008). which constricts sampled species to their respective clades to create distribution trees (Jetz et al., 2014). From the 9,000 distribution trees generated, we created a consensus phylogeny. Specific avian epithets listed follow Howard and Moore (2013). Species were then grouped into taxonomic clades for descriptive purposes, by order, family, or a combination of both where necessary.

The Cyt *b* haemosporidian Bayesian phylogeny was constructed in MrBayes 3.2. We selected the most appropriate model of nucleotide substitution as GTR+I model, as determined by both jModelTest 2.1. (Guindon et al., 2003; Darriba et al., 2012) and PartitionFinder 2.1.1 (Lanfear et al., 2017), and ran 10 million generations sampling every 1000 generations. A 20% percent burn-in of trees was discarded before creating a majority rule consensus tree. An outgroup was not specified and the final tree was rooted to the *Leucocytozoon* clade.

III.2.6 Prevalence Heat Map Analysis

Heat maps were created in R 3.3.2 (R Core Team, 2016) using the Superheat package (Barter and Yu). For the avian host taxonomy, we clustered species phylogenetically, and associated this with the prevalence of haemosporidian lineages (unscaled data), also clustered phylogenetically. We included duplicates of each

haemosporidian lineage if recovered from multiple localities, to characterize associations with avian hosts across localities.

We used the reduced data set of associated bioclimatic variables (Table 4, n=37) resulting from the PLSR analysis as bioclimatic input for heat map association with phylogenetically clustered haemosporidian lineages (based on the MrBayes analysis). The bioclimatic data were scaled from zero to one using a quantile preserving scale. We then used a hierarchical clustering analysis with pairwise Euclidean distances and a Ward's algorithm for the linkage method. Our haemosporidian lineages included duplicates of each lineage only if recovered from multiple localities, to characterize all environmental associations of each lineage.

III.3. Results

III.3.1 Parasite Lineage Diversity

We recovered a total of 85 unique haplotypes (hereafter referred to as lineages); 40 (47.1%) of these lineages were previously recorded in MalAvi/Genbank, while 45 (52.9%) are novel lineages. Of the 85 unique lineages, 41 (48.2%) were *Haemoproteus* haplotypes (28 of which are novel lineages), 33 (38.8%) were *Plasmodium* haplotypes (11 of which are novel), and 11 (13%) were *Leucocytozoon* haplotypes (six of which are novel lineages). For *Haemoproteus*, 28 of the 41 (70%) recovered lineages were novel, 11 of 33 (33.3%) were novel for *Plasmodium*, and six of 11 (54.5%) were novel for *Leucocytozoon* (Table 5).

Table 5. Haemosporidian detections across genera. Including novel and previously recovered lineages.

	Haemoproteus	*Leucocytozoon*	*Plasmodium*	Total
Individuals infected, % infected	48, 21.62%	14, 6.31%	74, 33.33%	85
Novel lineages	28	6	11	45
Previously recovered lineages	13	5	22	40

Pairwise genetic distance within each genus varied, *Haemoproteus* reached 10.9%, *Plasmodium* reached 10.1%, and *Leucocytozoon* reached 7.3% (Figure 6). Between group mean distance across genera was 9.6% for *Haemoproteus* and *Plasmodium*, 14.9% between *Haemoproteus* and *Leucocytozoon*, and 15.5% between *Plasmodium* and *Leucocytozoon*

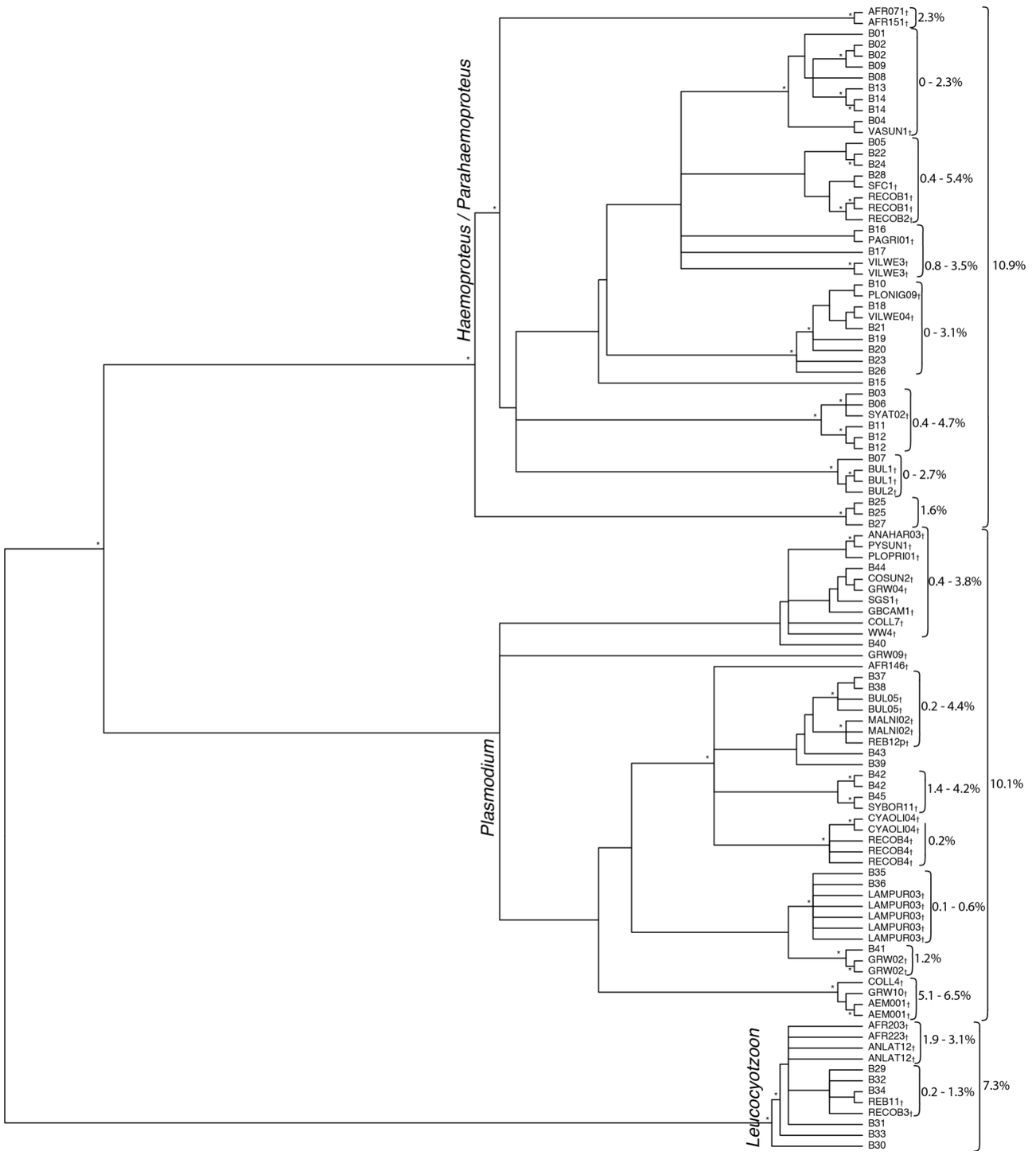

Figure 6. ***Cyt b*** **haemosporidian phylogeny for unique haplotypes across Benin localities. Bayesian analysis with posterior probability ≥ 95 notated with an *. Pairwise sequence divergence for intra-generic clades (range of % p-distance) and for each genus (maximum % p-distance) given.**

Co-infections, PCR recoveries of more than one genus of haemosporidian (*Haemoproteus*, *Leucocytozoon*, and/or *Plasmodium*) were recovered for 23 individuals (Table 6). *Haemoproteus* and *Leucocytozoon* co-infections were recovered in three individuals. *Haemoproteus* and *Plasmodium* co-infections were recovered across 14 individuals. *Leucocytozoon* and *Plasmodium* co-infections were recovered in six individuals. Co-infections of more than one lineage of the same genus in an individual were recovered in 21 individuals. Co-infections of *Haemoproteus* lineages were recovered in seven individuals. Co-infections of *Plasmodium* were recovered in 14 individuals (Table 6).

Table 6. Resolved co-infection matrix among and within haemosporidian genera.

	Haemoproteus	*Leucocytozoon*	*Plasmodium*
Haemoproteus	7	-	-
Leucocytozoon	3	0	-
Plasmodium	14	6	14

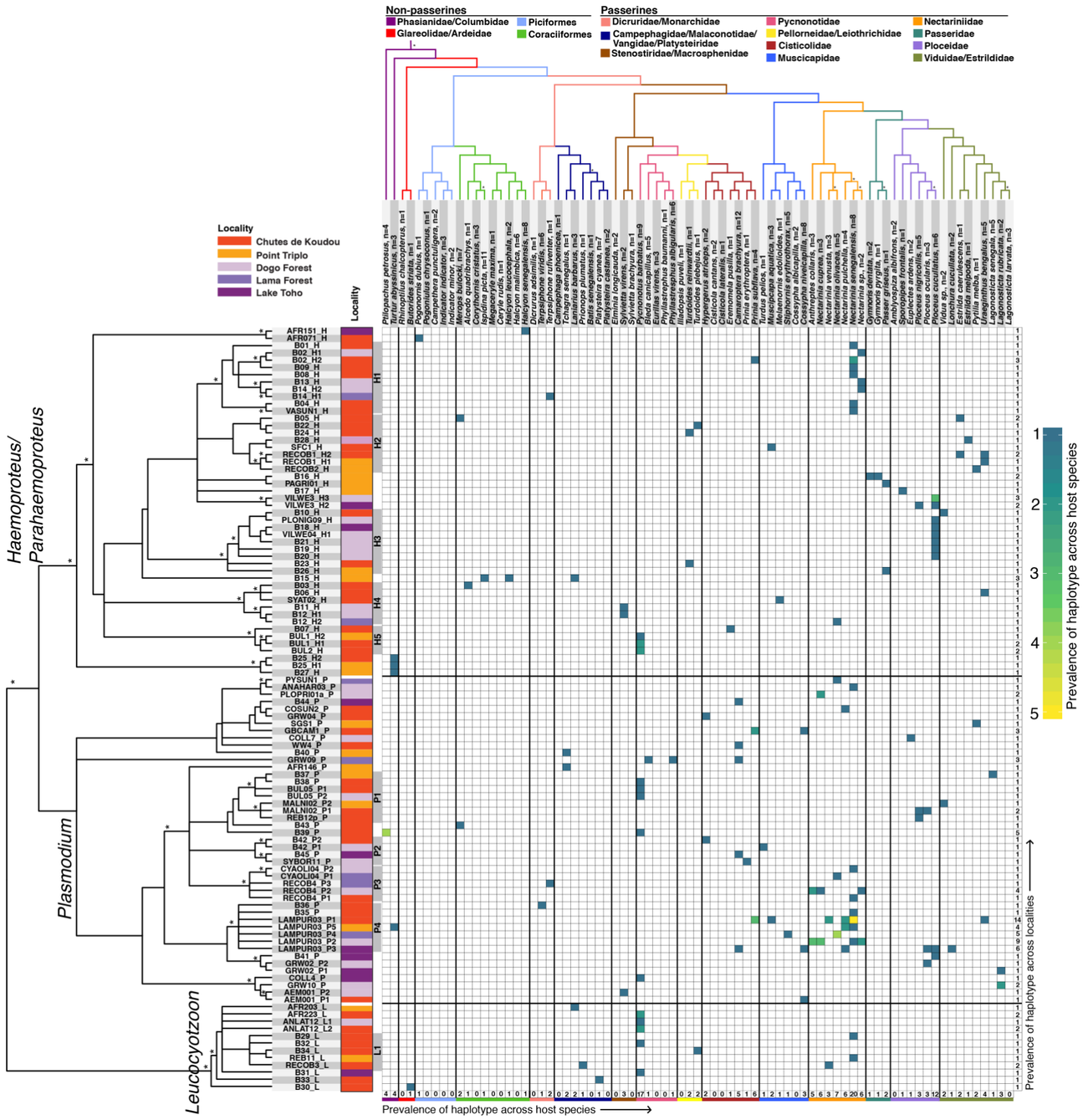

Figure 7. Haemosporidian phylogeny and heat map prevalence across avian host phylogeny. The heat map shows haemosporidian lineages recovered from Benin (left, shown in phylogenetic order) with prevalence as indicated by color gradient (Prevalence scale = 1-5, zeroes noted with a white square) across avian host taxonomy. Prevalence totals for each species sampled shown at bottom (ranging from 0 - 20). Prevalence of each lineage (across recovered localities), shown on right side (ranging from 1 - 14).

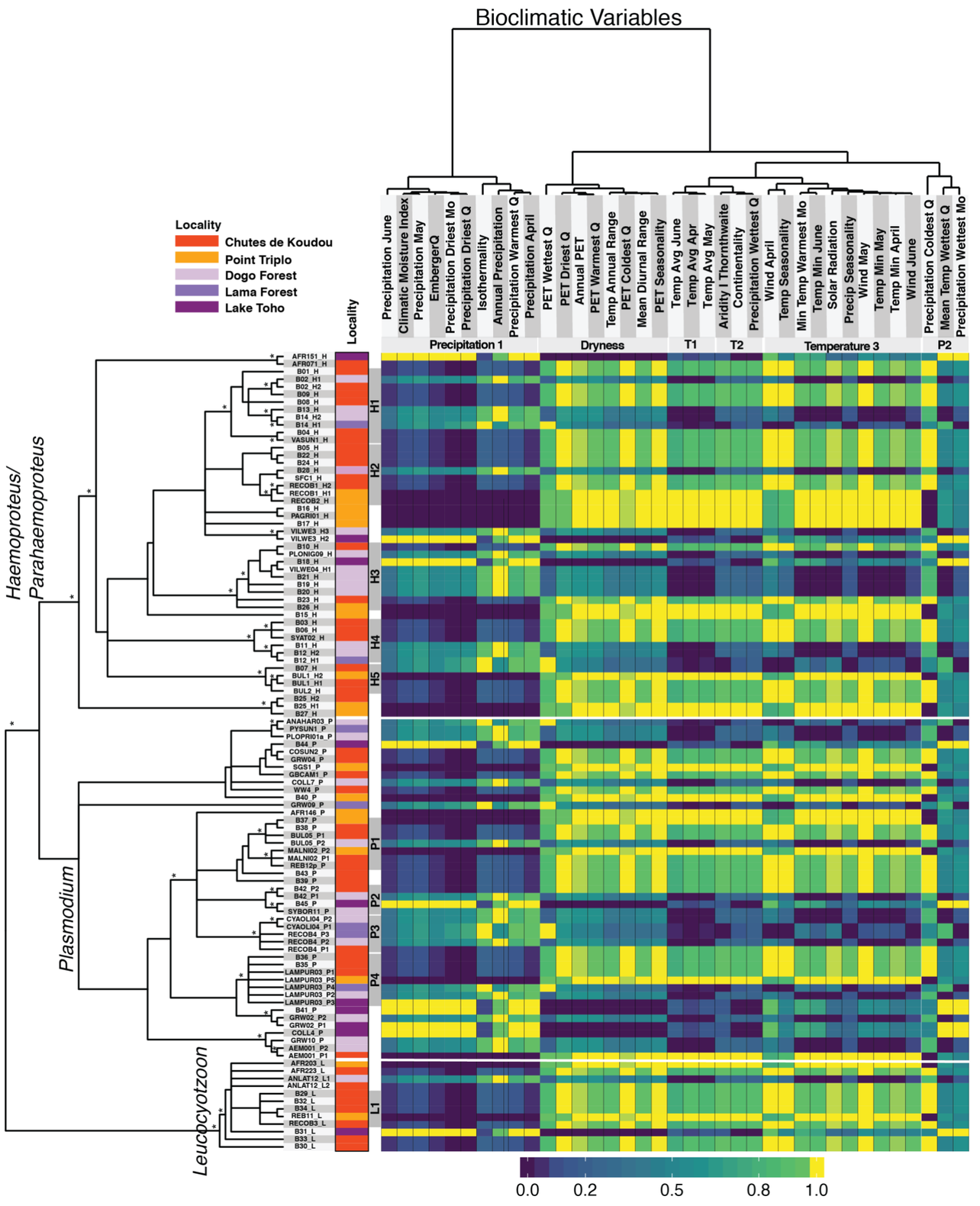

Figure 8. Dendrogram of hierarchically clustered bioclimatic variables (WorldClim V2 and ENVIREM) using Euclidean distances and Wards linkage method. Benin bioclimatic variables are scaled from zero to one using a quantile preserving scale. Dendrogram associated with parasite phylogeny showing unique lineages across localities (if recovered from multiple localities then lineage is listed for each).

We recovered 105 lineages which were unique across different sampling localities, meaning that 20 lineages that were recovered from multiple sampling localities (two to five localities) and are included in the analysis to characterize environments and host associations from recovered localities (Figs. 7 and 8).

The most highly recovered haemosporidian lineage LAMPUR03 (within clade P4), previously recorded in MalAvi, was here recovered a total of 38 times across 17 host species, across six avian families, and across all five localities (Figs. 7 and 8). The majority of LAMPUR03 host associations from our sampling were recovered in Nectariniidae (n = 26). The lineage LAMPUR03, as known from four previous MalAvi recoveries is widespread, recovered once from Gabon for a Sturnidae species, once in Bulgaria, and twice in Sweden. The three latter recoveries were from a European-African migrant *Ficedula* species. Lineage GRW09 is the most highly recovered MalAvi lineage across Africa (77 recorded MalAvi recoveries across eight countries), yet in this study it was only recovered three times, twice in Pycnonotidae and once in Cisticolidae, all from the Lama Forest locality (Fig. 7).

III.3.2 Parasite Host Associations

Avian sampling was diverse and includes representative of seven orders, 27 families, and 77 species, with most sampling consisting of Passeriformes (77%) (Table 7, Figure 9). We recovered positive PCR amplifications for haemosporidians from 113 of the total 222 individual avian hosts sampled (50.9%) (Table 7). We detected *Haemoproteus* in 48 of 222 individual hosts sampled (21.6%), *Plasmodium* in 74 of 222 individual hosts (33.3%), and *Leucocytozoon* in 14 of 222 individual hosts (6.3%).

Figure 9. Avian consensus tree for Benin sampling constructed using birdtree. Resolved from 9,000 distribution trees using the Hackett et al. (Hackett et al., 2008) as the backbone phylogeny. Avian species colored by clades which included order or family for descriptive purposes (Posterior probability ≤ 95 notated with an *).

We recovered associations for haemosporidian lineages across all avian taxonomic clades including non-passerines and passerines (Table 7). All orders, with the exception of Charadriiformes (nested within the Glareolidae/Ardeidae clade (grouped for description purposes), were positive for haemosporidian detection. Of the families sampled (n=27), 20 of 27 had one or more haemosporidian recoveries; the remaining seven families resulted in no recoveries, including one family within Charadriiformes, two families within Piciformes, and four families within Passeriformes (Figure 7, Table 7). Overall, of the 77 total avian species sampled, 55 species were positive for haemosporidian detection. *Haemoproteus* was recovered in all avian clades except for Glareolidae/Ardeidae. *Plasmodium* was recovered in all but four avian clades (Glareolidae/Ardeidae, Piciformes, Pellorneidae/Leiothrichidae, Passeridae). *Leucocytozoon* was taxonomically restricted to associations within the Glareolidae/Ardeidae, Campephagidae, Pycnonotidae, Pellorneidae/Leiothrichidae, and Nectariniidae clades.

Table 7. Host species taxonomic information including number of individuals sampled (N), localities sampled (C, P, D, L, T, A), frequency of detection across each genus *Haemoproteus* (H), *Leucocytozoon* (L), *Plasmodium* (P), ecoregion (E) and detections.

Order	Family	Species	Locality	N	H	L	P	E	Novel Lineage	MalAvi Lineage
Galliformes	Phasianidae	*Ptilopachus petrosus*	C	4	0	0	4	W	B39	
Columbiformes	Columbidae	*Turtur abyssinicus*	C, P (2)	3	3	0	1	W	B25, B27	LAMPUR03
Charadriiformes	Glareolidae	*Rhinoptilus chalcopterus*	C	1	0	0	0	W		
	Jacanidae	*Actophilornis africanus*	A	1	0	0	0	G		
Pelecaniformes	Ardeidae	*Butorides striata*	C	1	0	1	0	W	B30	
Coraciiformes	Alcedinidae	*Alcedo quadribrachys*	C	1	1	0	0	G	B03	
		Ceryle rudis	T	1	0	0	0	G/W		
		Corythornis cristatus	C, T	2	0	0	0	W		
		Halcyon leucocephala	C, P	2	1	0	0	W	B15	
		Halcyon malimbica	C (4), P (2)	6	0	0	0	G		
		Halcyon senegalensis	T	1	1	0	0	G/W		AFR151
		Ispidina picta	C (5), P (4), D, T	11	1	0	0	W	B15	
		Megaceryle maxima	C	1	0	0	0	W		
	Meropidae	*Merops bulocki*	C (5), P (2)	7	1	0	1	W	B05, B43	
Piciformes	Indicatoridae	*Indicator indicator*	C	3	0	0	0	G/W		
		Indicator minor	C	2	0	0	0	G/W		
	Picidae	*Campethera punctuligera*	C	2	0	0	0	W		
	Ramphastidae	*Pogoniulus chrysoconus*	C	1	0	0	0	G		
		Pogonornis dubius	C	1	1	0	0	W		AFR071
Passeriformes	Campephagidae	*Campephaga phoenicea*	P	1	0	0	0	W		
	Cisticolidae	*Camaroptera brachyura*	C (4), P (2), T (2), D (3), L	12	0	0	6	W	B44, B45	GBCAM1, GRW09, LAMPUR03, WW4
		Cisticola cantans	C	2	0	0	0	G		
		Cisticola lateralis	D	1	0	0	0	W		

Table 7. Continued

Order	Family	Species	Locality	N	H	L	P	E	Novel Lineage	MalAvi Lineage
Passeriformes	Cisticolidae	*Eremomela pusilla*	C	1	1	0	0	G	B07	
		Hypergerus atriceps	C	2	0	0	2	G	B42	GRW04
		Prinia subflava	C	4	1	0	6	W	B02	GBCAM1, LAMPUR03
	Dicruridae	*Dicrurus adsimilis*	C	1	0	0	0	W		
	Estrildidae	*Estrilda caerulescens*	C	1	2	0	0	G	B05	RECOB1
		Estrilda melpoda	D	1	1	0	0	W	B28	
		Lagonosticta larvata	C	3	0	0	0	W		
		Lagonosticta rubricata	D	2	0	0	3	W		GRW02, GRW10
		Lagonosticta senegala	C, P (4)	5	0	0	1	W	B37	
		Pytalia melba	P	1	1	0	1	G		RECOB2, SGS1
		Spermestes cucullata	T, D	2	0	0	1	W		LAMPUR03
		Uraeginthus bengalus	C (3) , P (2)	5	3	0	1	W	B06	LAMPUR03, RECOB1
	Leiothrichidae	*Turdoides plebejus*	C	1		1	0	W	B22, B34	
		Turdoides reinwardtii	C	1	2	0	0	G/W	B23, B24	
	Malaconotidae	*Laniarius barbarus*	C (2), P	3	1	1	0	W	B15	AFR203
		Tchagra senegalus	P	1	0	0	2	W	B40	AFR146
	Monarchidae	*Terpsiphone rufiventer*	L	1	1	0	1	G	B14	RECOB4
		Terpsiphone viridis	C (4), P (2)	6	0	0	1	G/W	B36	
	Muscicapidae	*Cossypha albicapilla*	C (2)	2	0	0	0	G		
		Cossypha niveicapilla	C (6), T (2)	8	0	0	3	W		AEM001, GBCAM1, LAMPUR03
		Melaenornis edolioides	C	1	1	0	0	G		SYAT02
		Muscicapa aquatica	C	2	1	0	1	G/W		LAMPUR03, SFC1
		Stiphornis erythrothorax	L	5	0	0	1	G		LAMPUR03, SFC1
	Nectariniidae	*Chalcomitra senegalensis*	C (6), P, D	8	7	2	11	G	B01, B02, B04, B08, B09, B29, B35	ANAHAR03, CYAOLI04, LAMPUR03, REB11, RECOB4, VASUN1
		Cinnyris cupreus	D	3	0	0	6	G		LAMPUR03, PLOPRI01, RECOB4

Table 7. Continued

Order	Family	Species	Locality	N	H	L	P	E	Novel Lineage	MalAvi Lineage
Passeriformes	Nectariniidae	*Cinnyris pulchellus*	C (2), P (2)	4	0	0	5	W		COSUN2, LAMPUR03
		Cinnyris venustus	C (2), T	3	0	1	3	G		LAMPUR03, RECOB3
		Cyanomitra olivacea	L	5	1	0	6	W	B12	CYAOLI04, LAMPUR03, PYSUN1
		Cyanomitra sp.	D	1	1	0	1	W	B02	LAMPUR03, RECOB3
		Hedydipna collaris	D	3	0	0	5	W		LAMPUR03, RECOB4
		Nectarina sp.	D	2	2	0	1	W	B13, B14	LAMPUR03
		Prinia erythroptera	D	1	0	0	1	G		SYBOR11
	Passeridae	*Gymnoris dentata*	P	1	1	0	0	W	B16	
		Gymnoris pyrgita	P	1	1	0	0	G	B16	
		Passer griseus	P	1	1	0	1	G	B26	PAGRI01
	Pellorneidae	*Illadopsis puveli*	L	1	0	0	0	G/W		
	Platysteiridae	*Batis senegalensis*	C	1	0	0	0	G/W		
		Dyaphorophyia castanea	L	2	0	0	0	G/W		
		Platysteira cyanea	C (6), L	7	0	1	0	G/W	B33	
	Ploceidae	*Amblyospiza albifrons*	T	2	0	0	0	G		
		Euplectes ardens	D	2	0	0	1	G	COLL7	
		Ploceus cucullatus	T (2), D (4)	6	10	0	2	G	B18, B19, B20, B21, B41	LAMPUR03, PLONIG09, VILWE04, VILWE3
		Ploceus nigricollis	C (2), T (3)	5	1	0	2	G/W		MALNI02, REB12p, VILWE3
		Ploceus ocularis	C, T (2)	3	0	0	3	W		GRW02, LAMPUR03, MALNI02
		Sporopipes frontalis	T	1	1	0	0	G	B17	
	Pycnonotidae	*Andropadus virens*	L (2), D	3	0	0	0	G		
		Bleda canicapillus	L	5	0	0	1	G		GRW09
		Phyllastrephus albigularis	L	6	0	0	1	G		GRW09

Table 7. Continued

Order	Family	Species	Locality	N	H	L	P	E	Novel Lineage	MalAvi Lineage
Passeriformes	Pycnonotidae	*Phyllastrephus baumanni*	L	1	0	0	0	W		
		Pycnonotus barbatus	C (6), P, T, D	9	5	7	4	G/W	B31, B32, B38	AFR223, ANLAT12, BUL05, BUL1, BUL2, COLL4
	Stenostiridae	*Elminia longicauda*	C	2	0	0	0	W		
	Sylviidae	*Sylvietta brachyura*	D	1	0	0	0	W		
		Sylvietta virens	D	2	2	0	1	W	B11, B12	AEM001
	Turdidae	*Turdus pelios*	D	1	0	0	1	W	B42	
	Vangidae	*Prionops plumatus*	C	1	0	1	0	W		RECOB3
	Viduidae	*Vidua sp.*	P, D	2	1	0	1	W	B10	MALNI02

We recovered several haemosporidian clades that were highly associated with avian host clades (Figure 7). Nectariniidae had the highest haemosporidian prevalence, with a total of 53 lineages (*Haemoproteus* = 11, *Plasmodium* = 39, *Leucocytozoon* = 3) recovered across 28 individuals. Nectariinidae exhibited three primary haemosporidian clades, one in *Haemoproteus* (H1), which was recovered from one northern (western Sudanian savanna) and two southern (Guinean forest-savanna) localities, and two *Plasmodium* clades (P3 and P4), which were recovered across all five localities. Clade P4 had numerous recoveries from other taxonomic host groups (Cisticolidae, Muscipadae, and Monarchidae) (Figure 7). Other strong associations were found in Pycnonotidae, with 19 lineages (*Haemoproteus* = 5, *Plasmodium* = 7, *Leucocytozoon* = 7) recovered from 24 individuals, which were recovered from four of five localities. The majority of Pycnonotidae infections were recovered from a single host species (*Pycnonotus barbatus)* which is one of the most common and widespread species across sub-Saharan Africa (Fishpool and Tobias, 2017). Pycnonotidae exhibited a predominant association with parasite the *Plasmodium* clade (P1), where all lineages were recovered from *P. barbatus* and from both northern localities and a single southern locality (the northernmost Dogo Forest locality). Cisticolidae had 15 lineages (*Haemoproteus* = 5, *Plasmodium* = 13) recovered from 23 individuals sampled and was found across four of five localities. Ploceidae had 20 lineages (*Haemoproteus* = 12, *Plasmodium* = 8) recovered from 19 individuals which were found across all localities with a distinctive *Haemoproteus* association (H3) which was found with detections across two of the southern localities (Dogo Forest and Lake Toho). The Viduidae/ Estrildidae avian host

clade recovered 16 lineages from 22 individuals (*Haemoproteus* = 8, *Plasmodium* = 8). The Estrildidae/Viduidae clade exhibited a discrete clade in *Haemoproteus* (H2) found in primarily in northern sites as well as the one recovery in the most northern southern locality, the Dogo Forest. *Leucocytozoon* L1, was detected in four passerine families, Vangidae, Pycnonotidae, Leiothrichidae, and Nectariniidae.

III.3.3 Environmental Associations

Hierarchical clustering of environmental variables resulted in five main clusters, where the variance in the variables is most similar. We categorize these five clusters as: 1) precipitation 1, 2) dryness (which includes primarily PET variables that measure the amount of evapotranspiration that would occur if water was available), 3) T1 and T2 (temperature related variables), and 4) temperature 3, 5) P2 (precipitation variables) (Figure 8; Table 4). We recovered varying degrees of clustering within the haemosporidian lineages across localities and environmental variables. We recovered *Haemoproteus* and *Plasmodium* from all five sampling localities. *Leucocytozoon* was recovered from four of five sampling localities, with none recovered from the Lama Forest. We observed higher associations of clustering of intra-generic clades in *Haemoproteus* (H2, H4 and H5) as compared to *Plasmodium* (P2, P3, P4), where clades cluster in bioclimatically similar localities. In *Haemoproteus* clade H3, clustering of variables is for high precipitation and lower temperature and aridity measures. Whereas for *Plasmodium* clade P1, we see the opposite, clustering is for low precipitation and high temperature and aridity measures (Figure 8).

For *Haemoproteus* richness and prevalence, the environmental predictor variable which most effectively maximized the covariance was precipitation of the driest quarter (PDQ; Table 4). For *Leucocytozoon*, we also recovered PDQ as the most informative predictor variable for richness, yet precipitation of the wettest month was most informative for prevalence. The most informative environmental predictor variables for *Plasmodium* were precipitation of the coldest quarter and PDQ. Our first clade in *Haemoproteus* H2 is most clearly clustered into two groups by variables in the aridity cluster, whereas clade H3 is most clearly differentiated by extremes in the variables for precipitation 1 and P2 (Figure 8). We see lower fidelity to locality and therefore bioclimatic clusters across *Plasmodium*. We recover mixing of climate and thus localities within clades P1, P2, and P4, with three to five localities represented within each clade. For *Leucocytozoon* our sole clade was highly clustered across bioclimatic variable and only recovered in both northern localities (Figure 8).

III.4. Discussion

In this study, we addressed both the importance of host associations and environmental variables in the structuring of haemosporidian lineages to provide insights into the ways in which climate affects the distribution of parasites and their associated hosts. We recovered substantial number of previously documented lineages (47.1%) in our sampling, which serves to highlight the degree of connectivity that exists across African regions, as well as migration aided transmission along the Eurasian-African flyway. Importantly however, we also recovered equally substantial new haemosporidian diversity as represented by 52.9% novel haplotypes. The high number of novel

haplotypes was not unexpected, given 1) that Benin has never been sampled for avian haemosporidians, and 2) that other recent parasite studies (e.g. on avian lice) have also shown newly recovered African species (Takano et al., 2017). We also determined that in our Benin sampling, bioclimatic variables do have an impact on haemosporidian genera richness and prevalence, although this effect varies across genera. Results suggest that these characterizations are genus specific but instead intra-generic clades as these may represent evolutionary units, though this determination was not the objective of this study.

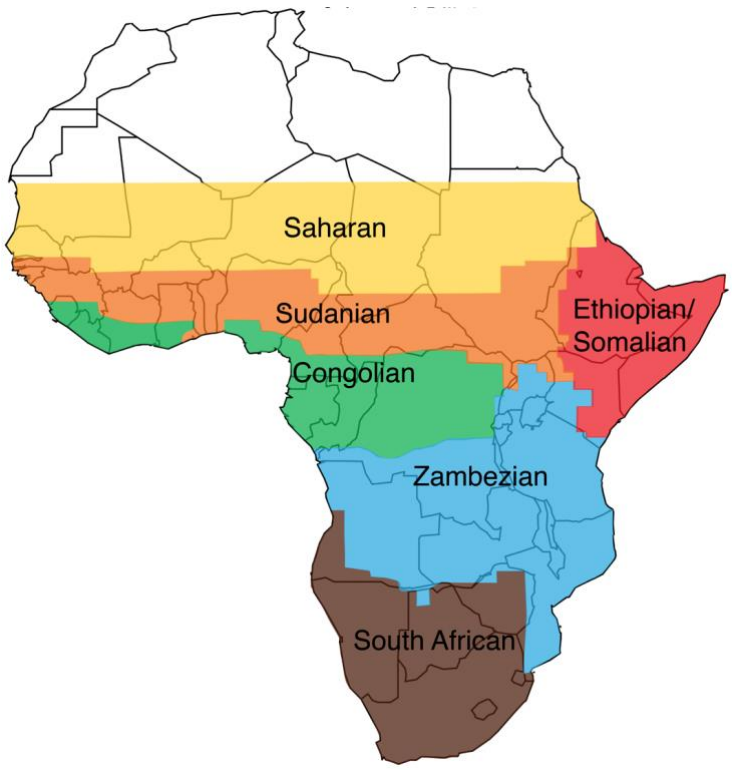

Figure 10. Bioregions as determined from clustering of avian diversity across Africa. (Adapted from Outlaw et al. 2017).

This is the first study addressing avian haemosporidians host and climate associations in the country of Benin. Sampling thus far remains low across the Sudanian biogeographic region (Outlaw et al., 2017), in which Benin is found in, with sampling in only a few Sudanian countries (Figure 10). Nigeria has the highest haemosporidian sampling (n = 201 MalAvi lineages); these samples are from both bioregions in which Nigeria is found in, Sudanian and Saharan (Dinerstein et al., 2017). A handful of other samples have been taken from other Sudanian bioregion countries (all from the Eastern Guinean Forest), for example Ghana (n = 38 MalAvi lineages) and Senegal (n = 2 MalAvi lineages). We recovered higher infections of *Haemoproteus* as compared to a Sudanian bioregion meta-analysis of Malavi lineages (36.5% based on 167 lineages total Sudanian lineages; Outlaw et al. 2017); however, given the high world-wide diversity of biting midge species (Culicoides) this is not unexpected and results may have been due to limited sampling (Mellor et al., 2000). We recovered similar *Plasmodium* and lower *Leucocytozoon* infections as compared to the broader Sudanian bioregion (Outlaw et al. (2017). We believe that our low recovery of *Leucocytozoon* is due to the general lack of streams and bodies of water typically used by black flies near many of our sampling sites, which would lower black fly vector abundance (Sutcliffe, 1986). While black flies are indeed present in Benin (current species inventories are reflective of medically relevant species primarily relating to cattle), vector diversity and abundance has not been well assessed across Africa (Adler and Crosskey, 2013).

Host association and bioclimatic variables are likely critical factors structuring the distribution of haemosporidian lineages, and our analyses are an attempt to visually

represent these relationships (Figures 8, 9). The relationships recovered reflect a wide spectrum of lineage clusters associations, in which they are constrained to host taxonomy, by bioclimatic characteristics or both. For example, clade H1 within *Haemoproteus* is strongly associated with the avian family Nectariinidae, with only two recoveries found outside this clade. Climatically, H1 is found in two of the southern sampling localities (Dogo Forest, Lama Forest), with multiple recoveries from only one of the northern localities (Chutes de Koudou). Chutes de Koudou is the most climatically similar to southern sites with respect to variables within Precipitation 1. So, while H1 is generally constrained by host taxonomy in our sampling, it is also structured geographically by precipitation, indicating vector constraint tied to precipitation. Alternatively, clade H2 lineages are recovered across four diverse host taxonomic groups and primarily recovered from climatically similar northern localities; the exception was a recovery from the Dogo Forest. Therefore, clade H2 distribution is apparently more impacted by climate variables than by host association and indicates broadly distributed vectors. This is supported by the fact that previously known haplotypes in this clade (MalAvi lineages SFC1, RECOB1, RECOB2) are recorded as occurring in the same families recovered in this study. In *Plasmodium,* we see less host specificity for lineage associations as all clades are recovered broadly across two to six taxonomic groups. We see varying levels of association across climatic variables, with lineage clusters tied to as few as 3 localities (P1, P2, and P3) and as many as all 5 (P4). The patterns of association for clades that cluster strongly for environmental extremes (such as clades H3 and P1; Figure 8) indicate habitat specificity of the vectors.

While these visualizations provide insight into limitations of the distribution of phylogenetic sub-clades of haemosporidian lineages, we know there are other factors that shape distributions that are not represented here. These include, but are not limited to, competent host availability as well as non-competent host species densities (Keesing et al., 2006). Parasite transmission can be reduced if non-host density exceeds competent host density. Further, abiotic factors, which are difficult to measure, include water accumulation and distance to bodies of water. Given that the requirements of precipitation and standing water vary greatly haemosporidian vector groups (from puddles of water to running streams) it is difficult to capture the broad variation needed to determine the distribution of vector groups and their subsequent presence in host populations. While the broad associations of *Plasmodium* lineages across sampled avian host taxonomy were as expected, the high parasitism rates associated with some taxonomic groups (Figure 7; *Haemoproteus*: H3 in Ploceidae, H2 in Estrildidae, and *Plasmodium*: P1 in Nectariniidae) unexpected. Confirmation of this would require further quantification of parasitemia (real-time quantification PCR) as it is difficult for histology and microscopy to detect low or inactive infections (Valkiunas, 2005).

Disentangling distributional patterns in the avian haemosporidian system is complex and affected by several factors, such as vector associations and bioclimatic constraints across ecoregions. We have limited knowledge of vector distributions and abundances, and bioclimatic data in this study were selected to take into account vector/parasite development requirements. Increasing the understanding of the ecological constraints of the various vector groups associated with each genus is a

priority to inform parasite distribution modeling. Although we sampled across just two discrete ecoregions, our results indicate the importance of assessing bioclimatic variables when assessing haemosporidian distributions across regions and habitats. As such, we believe these types of across-ecoregion comparisons, which also include phylogenetic analyses, will be increasingly informative at a greater geographic scale, particularly with increased replicates across similar habitats. Future work should expand to examine haemosporidian diversity across Africa's 110 diverse ecoregions (Figure 4), the diversity of these ecoregions is well exemplified in South Africa's 18 ecoregions. Increasing sampling localities will further allow the implementation of predictive modeling. To determine the effect of the host and bioclimatic characteristics, we need to assess the patterns at a larger geographic scale encompassing multiple ecoregions and diverse habitats.

CHAPTER IV

HAEMOSPORIDIAN DIVESITY AND AVIAN HOST ASSOCIATION RECOVERED FROM THE DEMOCRATIC REPUBLIC OF THE CONGO

IV.1 Introduction

Understanding the diversity and distribution of parasites and pathogens is becoming exceedingly important due to increases in habitat and climate alterations, as well as the effects of introduced species. Further, parasites are known to impact community structure and evolutionary trajectories (Ricklefs 2010). Migratory transmission has been documented for avian malaria, West Nile Virus, and avian influenza virus, among others (van Riper III *et al.* 1986; Randolph & Rogers 2010; Van Hemert, Pearce & Handel 2014). For example, haemosporidians (protozoan vector transmitted parasites) can affect host fitness from physical condition and reproductive success to reduced survival rates; they can also be a source of emerging disease in non-endemic ranges (van Riper III *et al.* 1986; Kilpatrick *et al.* 2006b; Knowles, Palinauskas & Sheldon 2010). Therefore, knowledge of diversity and distributional patterns are necessary to detect shifts in distributional ranges of parasites and pathogens, and the causative effect of disease and parasitism on host species.

Avian haemosporidians are a highly diverse system, and as such are ideal for studying host-parasite interactions. Haemosporidian parasites range nearly world-wide (except Antarctica) and parasites have been found to infect much of avian tree of life. Avian haemosporidian genera include *Haemoproteus* (including *Parahaemoproteus*),

Leucocytozoon, and *Plasmodium*, and each genus is vectored by a different suite of dipterans. *Haemoproteus* is vectored by biting midges (Ceratopogonidae) and louse flies (Hippoboscidae), *Plasmodium* is vectored by mosquitos (Culicidae), and *Leucocytozoon* is vectored by black flies (Simuliidae) (Valkiunas 2005). In the haemosporidian life cycle, dipterans are the vectors and the definitive host (where sexual reproduction occurs) of haemosporidians, whereas vertebrates, such as birds, are intermediate hosts and the location of asexual reproduction.

Avian haemosporidian lineages have been found to range widely in their host associations, from host specific (restricted to one or few hosts) to host generalists (found across taxonomically diverse hosts) (Ricklefs & Fallon 2002; Beadell *et al.* 2004, 2009; Poulin 2007; Hellgren, Pérez-Tris & Bensch 2009; Drovetski *et al.* 2014; Ricklefs *et al.* 2014). Host specificity was previously believed to be a characteristic consistent at the haemosporidian species-level; however, current studies indicate that all genera exhibit both host specificity and host generalism and this range of infection across hosts is more conserved at the intra-specific level (Gager *et al.* 2008; Hellgren *et al.* 2009; Medeiros, Ellis & Ricklefs 2014). Host-parasite (haemosporidian) interactions are undoubtedly influenced by vector encounter rate with hosts, further vectors have been found to aid in host switching occurrence (Gager *et al.* 2008). Indicating that vector associations may inform parasite specificity and generalism. Parasite host specificity reduces the number of possible hosts; however, specificity has been found to increase the intensity of parasitism, thereby making specialists more proficient at infecting their hosts (i.e. fitness within the host) (Ward 1992; Medeiros *et al.* 2014).

There has been growth in the use of avian haemosporidians as a model system for vector borne parasite systems, with many studies seeking to inform host-parasite interactions (Ishtiaq *et al.* 2006; Drovetski *et al.* 2014; Lutz *et al.* 2015; Olsson-pons *et al.* 2015). In order to determine patterns of association and diversification for haemosporidians and their hosts, broad assessments of avian host communities are needed across diverse regions. Sub-Saharan Africa offers a diversity of habitats and ecoregions and as such houses some of the highest avian diversity in the World (Mittermeier *et al.* 1998; Brooks *et al.* 2001; Linder *et al.* 2012). Consequently, avian haemosporidian diversity is expected to be correspondingly diverse. The majority of sub-Saharan African haemosporidian studies have thus far assessed limited host species diversity (i.e. targeted sampling). While taxonomically diverse studies of avian haemosporidians do exist in sub-Saharan Africa (Beadell *et al.* 2009; Okanga *et al.* 2014; Lutz *et al.* 2015) these studies are geographically limited. In general then, there is a deficit of community level analyses across Africa, which limits our ability to extrapolate for comparative analyses and inform patterns at the broader scales (see Outlaw *et al.* 2017).

In this study, we seek to expand on community level assessments in Africa by assessing haemosporidian diversity from a suite of African lowland forest tropical birds sampled from the Democratic Republic of the Congo. We examine the avian community broadly, although sampling is biased towards understory species due to sampling methods used (described below). Given the lack of research on avian haemosporidians from this region, we expected to recover both a high number of novel lineages, as well

as new host associations. We examine patterns of parasite host specificity and generalism across parasite genus, and avian species. Further, we examine avian families for which we had dense sampling to determine whether parasite lineage constraint patterns were apparent and whether these are associated with host evolutionary traits which are correlated with vector occurrence. In other words, is there evidence of host specificity, and if so, can such associations be linked to avian life history traits such as behavioral interactions and foraging strata.

IV.2 Methods

IV.2.1. Sampling localities and avian sampling

Birds were sampled in the Democratic Republic of the Congo in August of 2009 and 2010 as part of a broader biodiversity study where 421 individual birds were sampled (Table 8). Sampling took place north and south of the Congo River near Kisangani (Fig. 1). Two localities (Masako and Yoko Forest Reserves) where sampled in 2009, and two (Yelenge and Djabir) were sampled in 2010. All sampling localities are within 30 km from the Congo River (Figure 11)

Table 8. Host species taxonomic information including localities sampled (Y, M, D, YE), number of individuals sampled (N), frequency of detection across each genus *Haemoproteus* (H), *Leucocytozoon* (L), *Plasmodium* (P), and novel/previous detection of lineages.

Order	Family	Species	Locality	N	H	L	P	Novel Lineage	MalAvi Lineage
Caprimulgiformes	Caprimulgidae	*Caprimulgus batesi*	Y	1	0	0	0		
Cuculiformes	Cuculidae	*Cercococcyx olivinus*	Y	1	0	0	0		
Columbiformes	Columbidae	*Turtur brehmeri*	Y	1	1	0	0	DRC18	
Bucerotiformes	Bucerotidae	*Tockus fasciatus*	Y	1	0	0	0		
Coraciiformes	Alcedinidae	*Alcedo quadribrachys*	M (3), Y	4	1	0	0		ALCQUA01
		Corythornis leucogaster	YE (2)	2	0	0	0		
		Halcyon badia	Y	1	0	0	0		
		Ispidina lecontei	D (2), M (4), YE (2), Y (6)	14	5	0	0		ALCLEU02, ALCLEU03, CELEC01
		Ispidina picta	M	1	0	0	0		
Piciformes	Ramphastidae	*Pogoniulus atroflavus*	M	1	0	0	0		
		Pogoniulus scolopaceus	D, YE (3)	4	0	0	0		
		Pogoniulus subsulphureus	D	1	0	0	1		MILANS06
	Indicatoridae	*Indicator maculatus*	D	1	0	0	2		ANAHAR03, RECOB4
	Picidae	*Campethera nivosa*	YE (4), M (2)	6	1	0	1		COLL2, GRW09
		Verreauxia africana	M	1	0	0	0		
Passeriformes	Calyptomenidae	*Smithornis rufolateralis*	D, M, Y	3	0	0	0		
	Platysteiridae	*Dyaphorophyia castanea*	M, Y	2	0	2	1	DRC6, DRC8, DRC48	
	Monarchidae	*Terpsiphone batesi*	D, M (3), YE (2), Y (3)	9	6	1	0	DRC15	TERUF01
		Terpsiphone rufiventer	D (2), M (2), YE (3), Y (5)	12	5	3	0	DRC7, DRC15	ARITEP02, TERUF01
	Turdidae	*Neocossyphus poensis*	D, M, YE (2), Y (5)	9	0	2	5	DRC16, DRC50	ACCTAC01, ANLAT04, ANLAT10, GRW09, PLACAS02
		Stizorhina fraseri	YE (4), Y	5	2	1	3	DRC4, DRC50	ACCTAC01, COLL2, PLACAS02
	Muscicapidae	*Alethe diademata*	D (3), M (3), YE (5), Y (7)	19	3	0	9	DRC19, DRC22, DRC42	ALEDIA02
		Muscicapa olivascens	YE	1	0	0	1		PLOPRI01
		Myioparus griseigularis	M	1	0	0	0		
		Stiphornis erythrothorax	M (2), YE (3), Y (9)	14	0	0	10	B50, DRC36, DRC51	PHICT01
	Nectariniidae	*Cyanomitra olivacea*	D (14), M (10), YE (7), Y (11)	42	3	7	28	DRC13, DRC15, DRC16, DRC17, DRC25, DRC27	ANLAT04, ANLAT12, CYAOLI04, CYAOLI05, CYAOLI09, PHICT01, PYSUN1, RECOB4
		Deleornis fraseri	D (6), Y	7	1	1	8	DRC16, DRC26, DRC28, DRC19	ANAHAR03, PYSUN1, RECOB4, RECOB4a, RECOB4b
	Ploceidae	*Malimbus malimbicus*	Y (2)	2	0	1	1	DRC5	CYAOLI09
	Estrildidae	*Nigrita bicolor*	M	1	0	0	0		
		Nigrita fusconotus	YE	1	0	0	0		
		Parmoptila rubrifrons	D	1	0	0	0		

Table 8. Continued

Order	Family	Species	Locality	N	H	L	P	Novel Lineage	MalAvi Lineage
Passeriformes	Estrildidae	*Parmoptila woodhousei*	Y	1	0	0	0		
		Pyrenestes ostrinus	M (2)	2	1	0	0	DRC24	
		Spermophaga poliogenys	YE (4), Y (3)	7	1	1	2	DRC43	ALCLEU02, ANLAT12, PHICT01
	Nicatoridae	*Nicator choloris*	Y	1	0	1	1	DRC14, DRC44	
	Scotocercidae	*Hylia prasina*	D (7), M (7), YE (9), Y (6)	29	2	4	2	DRC11, DRC32	ALCLEU02, HYLPRA01, PHICT01, TERUF01
	Pellorneidae	*Illadopsis albipectus*	M (6), YE, Y (3)	10	1	0	5	DRC22, DRC45, DRC46, DRC47, DRC53	
		Illadopsis fulvescens	M (8), YE (4)	12	0	1	4	DRC33, DRC34	ANLAT12
		Illadopsis rufipennis	M (3), YE (7), Y (7)	17	4	0	8	DRC20, DRC21, DRC31, DRC34, DRC35, DRC49	CYAOLI09, RBS4
	Pycnonotidae	*Baeopogon indicator*	YE	1	0	0	1		PHICT01
		Criniger calurus	D (2), M (4), YE (4), Y (6)	16	0	3	0	DRC10	ANLAT11
		Criniger chloronotus	Y (2)	2	0	2	2	DRC1, DRC10	PHICT01
		Eurillas latirostris	D (10), M (13), YE (10), Y (17)	51	0	4	10	DRC!2, DRC38	ANLAT04, ANLAT12, ARITEP02, PHICT01
		Eurillas virens	D, M (9)	10	0	1	5	DRC9, DRC37	CYAOLI02, CYAOLI04, PHICT01
		Ixonotus guttatus	YE (4)	4	0	1	2	DRC3, DRC30	
		Phyllastrephus albigularis	D (2), M (4)	8	0	0	1		ANLAT04
		Phyllastrephus icterinus	D (7), M (6), YE (4), Y (7)	27	6	2	5	DRC23, DRC40	ANLAT02, ANLAT04, ANLAT12, CYAOLI02, PHICT01
		Phyllastrephus xavieri	M (5), YE (8), Y (4)	17	1	1	5	DRC39, DRC53	ANLAT02, ANLAT04, ANLAT12, PHICT01
		Bleda eximius	M (5), YE (5), Y (6)	16	0	5	3	DRC1, DRC2, DRC52	ANLAT10, GRW09, LAMPUR03
		Bleda syndactylus	D (3), M (3), YE (4), Y (9)	19	2	1	7	DRC22, DRC40, DRC41, DRC52	ANLAT04, ANLAT11, PHICT01, RBS4

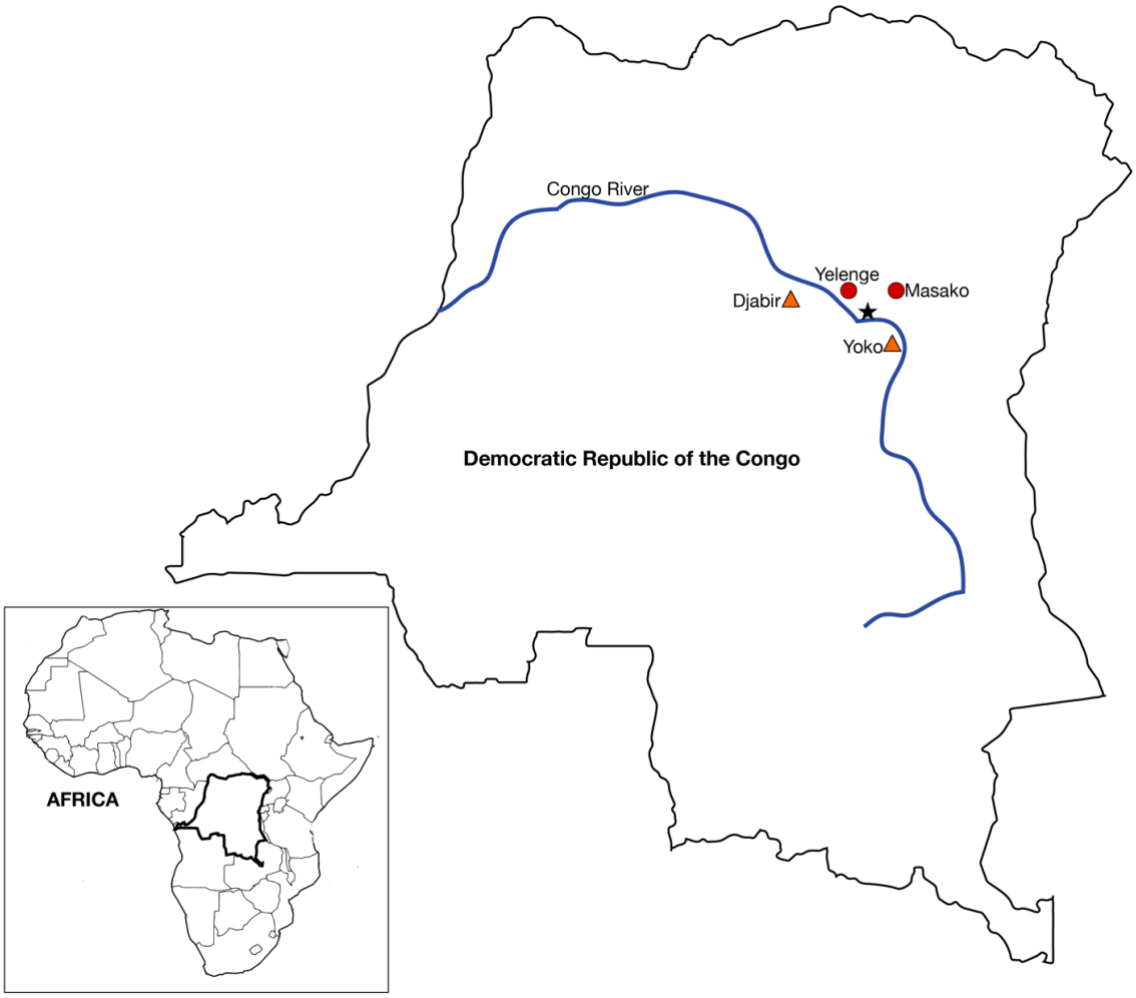

Figure 11. Map of the Democratic Republic of the Congo with sampling localities north and south of the Congo River and surrounding the city of Kisangani (black star). Localities of Yelenge and Masako, denoted by red circles, are north of the Congo River. Localities of Djabir and Yoko, denoted by orange triangles, are south of the Congo River.

For analysis, we screened blood samples for all individuals except 38 individuals where blood was not available and pectoral muscle was screened in its place. All voucher specimens collected are accessioned in the Biodiversity and Research Teaching Collections, at Texas A&M University. We collected all specimens under protocols

approved by the Institutional Animal Care and Use Committee at Texas A&M University.

IV.2.2 Molecular assessment of avian haemosporidians

We followed molecular protocols previously described in Harvey and Voelker (2017) for DNA extraction, PCR reactions and sequencing of positive PCR reactions. Briefly, we targeted a 505 base pair (bp) fragment of the mitochondrial (mtDNA) cytochrome *b* gene (*Cyt b*) using multiple primer pairs previously published in Drovetski et al. (2014). This fragment encompassed the entire 479 bp of the standard gene region collected in the MalAvi avian haemosporidian database (Bensch, Hellgren & Pérez-Tris 2009). Collectively, these primers amplify all three genera of avian haemosporidians: *Haemoproteus* (to include subgenera *Parahaemoproteus* and *Haemoproteus*), *Leucocytozoon*, and *Plasmodium.*

We extracted whole genomic DNA from avian blood and pectoral muscle samples using the E.Z.N.A. Tissue extraction kit (Omega Bio-Tek, Norcross, GA) and standard protocols except for the final elution was 75 µl to increase overall DNA yield. Polymerase chain reaction (PCR) was used to identify haemosporidian infection by amplifying a fragment of the mtDNA cyt*b* gene and multiple primer pairs were used to amplify across known avian haemosporidian genetic diversity. Three primer pairs consisting of the same forward primer (UNIVF) and one of three distinct reverse primers: UNIVR1, UNIVR2 or UNIVR3 were used, yielding 505, 535, and 571 base pairs (bp) (Drovetski et al., 2014), respectively. These primers encompass the entire 479 bp that are the standard region that is collected in the MalAvi avian haemosporidian

database (Bensch et al., 2009). Collectively, these primers amplify all three genera of avian haemosporidians: *Haemoproteus* (to include subgenera *Parahaemoproteus* and *Haemoproteus*), *Leucocytozoon*, and *Plasmodium*. PCR amplification was carried out in 18.75 µl reaction containing 1 x GoTaq Flexi buffer, 2.5 mM MgCl2, 0.2 mM of dNTP, 0.19 µl of each primer and 0.9375 µ of GoTaq Flexi DNA polymerase (Promega Madison, Wisconsin, USA) with 1 µl of DNA template. Each sample (both blood and pectoral muscle) was tested via PCR for each primer pair, and up to three times per primer pair. If a positive amplicon was not detected and successfully sequenced during this process, we considered the sample to negative for haemosporidians. All PCR's included four positive controls as well as 2 negative controls. Automated sequencing was performed bi-directionally using BigDye (Applied Biosystems, Foster City, CA, USA) and products were run out using an ABI 3730 at Beckman Coulter Genomics (Danvers, MA, USA).

Sequences were verified for quality base by base and aligned by eye using Geneious 6.1.8 (http://www.geneious.com, Kearse et al., 2012). Multiple infections were determined by the presence of multiple peaks on both chromatograms at one or more base positions (Harvey & Voelker 2017). Due to the high probability that sequences with less than three DNA positions displaying multiple peaks were sequencing errors and not true multiple infections (Szymanski & Lovette 2005), we treated these as single infections. After being verified with criteria for peak similarity, confidence score and visual assessment we processed double infection data (n = 8) with assigned IUPAC ambiguity codes. We then reconstructed single infection haplotypes (Browning &

Browning 2011) using Phase 2.1 (Stephens, Smith & Donnelly 2001) as implemented in DnaSP 25.10.1 (Librado & Rozas 2009) along with all similar genetic sequences from MalAvi, including all data that is ≥ 97% identity match.

All data were identified to genus by use of MalAvi blast (Version 2.2.8, Bensch et al., 2009) and NCBI BLAST (Altschul *et al.* 1990) functions. Criteria of one base pair was followed to differentiate genetic lineages (i.e. unique haplotype) (Bensch *et al.* 2000, 2004). Sequences were assigned the MalAvi lineage name (if identified as a complete match) or identified as novel detections (Genbank accession numbers "pending").

IV.2.3 Delimitation of parasite clades of interest

To assess diversity of haemosporidians we need to determine the biological units we are addressing. The delineations for haemosporidian evolutionary units (species and lineages), are poorly understood. Species descriptions have been primarily based on morphological descriptions of circulating red blood cell life stages (meronts and gametocytes) using microscopy of blood smears, while some species description are across all life stages (this is methodologically more challenging and therefore rare) (Valkiunas 2005). Currently there are just 220 morphologically described haemosporidian species (MalAvi 2.3.3). Molecular determinations, using the single nucleotide substitutions within the 479 bps of the *Cyt b* region, have yielded 2,876 molecular lineages thus far (MalAvi 2.3.3). In determining what constitutes a new molecular lineage, a one base pair difference has been accepted as the best practice

(Bensch *et al.* 2000, 2004) and is supported by the fact that Haemosporidian lineages have not shown evidence of recombination events (Joy 2003; Bensch *et al.* 2004, 2009).

Delimitations of species are more complex and are not standardized across haemosporidian taxa. The sequence divergence between the human *Plasmodium falciparum* and the chimpanzee *P. reichenowi* is 3.3%. Sequence divergence within *Plasmodium falciparum* parasites reaches 0.2% across all of *Cyt b*, consisting of six haplotypes with single base pair substitutions (Joy 2003). When examining the same *P. falciparum* data set restricted down to the 505 base pairs assessed in this study pairwise sequence divergence reaches 0.4% and consists of four haplotypes with single base pair substitutions. The globally distributed lineages of avian *P. relictum* demonstrate a 7.6% sequence divergence (Beadell *et al.* 2006). Avian *P. relictum* from the Hawaiian Islands was previously believed to show no sequence divergence (Beadell *et al.* 2006). A study by Jarvi et al. (2013) demonstrated a high number of variant haplotypes recovered from Hawaiian *P. relictum* using deep sequencing. However, these variants had low coverage and included non-synonymous substitutions resulting in stop codons, suggesting that these variants may be a result of sequencing error or random mutations which are biologically not meaningful; thus, supporting the genetic conservatism previously detected in Hawaiian *P. relictum*. This variability in the divergence rates across groups supports the argument to range criteria (1 -5 %) for species delimitations and use of other characteristics as support, including morphology where available, host association, and locality (Outlaw & Ricklefs 2014).

Here then, we are not attempting to determine species as morphological data is missing. Instead we use molecular lineages to reconstruct phylogenetic clades of haemosporidians to determine patterns of association across hosts and climate. Clades are selected with the criteria of being reciprocally monophyletic and having a within group pair wise sequence divergence of less than 5.5% though most are more conservative. Pairwise sequence divergence was measured using MEGA version 7.0.14 (Kumar, Stecher & Tamura 2016).

IV.2.4 Phylogenetic analysis

Avian host phylogenies were created, to determine taxonomic host breadth of haemosporidians, using birdtree.org, which uses relaxed molecular trees of well-supported clades along with a fossil calibrated backbone for constricted clades to create distribution trees (Jetz et al. 2012; Jetz et al. 2014). We selected a phylogeny subset of all avian species sampled in the Democratic Republic of the Congo using 10,000 trees from Hackett et al. (2008). Avian specific epithets and taxonomy and classification follow the Howard and Moore checklist (Dickinson & Remsen 2013).

A *Cyt b* haemosporidian Bayesian phylogeny was constructed in MrBayes 3.2. We selected the most appropriate model of nucleotide substitution as GTR+I model, as determined by both jModelTest 2.1. (Guindon, Gascuel & Rannala 2003; Darriba *et al.* 2012) and PartitionFinder 2.1.1 (Lanfear *et al.* 2017), and ran 10 million generations sampling every 1000 generations. A 20% percent burn-in of trees were discarded before creating a majority rule consensus tree. An outgroup was not specified and the final tree was rooted to the *Leucocytozoon* clade. We also constructed minimum-spanning

haplotype networks (Bandelt, Forster & Röhl 1999) using PopART (Leigh & Bryant). We did this only for avian species and families for which we had comparatively dense sampling, to attempt to characterize diversity patterns and associations of their haemosporidians. Species within the family network analysis were characterized for two life history behaviors: 1) grouping behavior (which include solitary, paired, or flocking behaviors) and 2) foraging strata (which describes the level of the forest where species forage or primarily reside), as described in del Hoyo et al. (2017). These life history behaviors may inform the host-vector encounter rate because 1) just as avian species differ in preference for foraging strata so do haemosporidian vector groups, and 2) avian grouping behavior may restrict (solitary species) or greatly increase (flocking species) the likelihood of those vector encounters.

IV.2.5 Prevalence Heat Map Analysis and Host Specificity Indices

Heat maps were created in R 3.3.2 (R Core Team, 2016) using the Superheat package (Barter & Yu) in order to visualize lineage presence across the parasite and host phylogenies. For the avian host taxonomy, we clustered species phylogenetically, and associated this with the prevalence of haemosporidian lineages, also clustered phylogenetically. The heat map data was prevalence, and is therefore unscaled data. Each haemosporidian lineage was assessed for all avian host associations recovered.

We examined the host breadth of each lineage using the standardized host specificity index S_{TD}* (Poulin & Mouillot 2005). The S_{TD}* metric determines host specificity using the taxonomic distance of hosts used by a parasite. This is measured as the number of steps required to reach a common ancestor (from genus, species, family,

order, superorder, to class), where more steps result in a higher S_{TD}* output suggesting low host specificity. The taxonomic distance is then weighted by the prevalence. Taxonomic S_{TD}* values of 1 are given for all lineages infecting more than one host species, while lineages infecting a single host species are automatically assigned an S_{TD}* value of 1.

We examined all lineages recovered that have been previously been recorded in MalAvi 2.2.8 (Bensch *et al.* 2009) in order to determine whether recovered specificity measures varied with added recoveries, given no previous sampling for haemosporidian diversity exists for this region of Africa. The metric S_{TD} was used, which measures the taxonomic distance of hosts without taking in to account prevalence across cross host species. S_{TD} was determined for data collected in DRC (local) and DRC plus all MalAvi records (global).

IV.3 Results

Avian sampling was diverse and included seven orders, 20 families, and 51 species with the majority of sampling consisting of Passeriformes (90.7%) (Figure 12; Table 8). Sampling density across species varies from one to 51 individuals sampled across species (Figure 12; Table 8). Of the 421 individuals sampled, we recovered 187 individuals with positive PCR detections for haemosporidians, resulting in an overall 42.4% infection rate. From our sequencing of these positive detections, we recovered a total of 84 unique haplotypes (hereafter referred to as lineages); 24 (28.6%) of these lineages were previously recorded in MalAvi/Genbank, while 60 (71.4%) are novel lineages. We detected *Haemoproteus* in 44 of 421 individuals (10.5%), *Leucocytozoon*

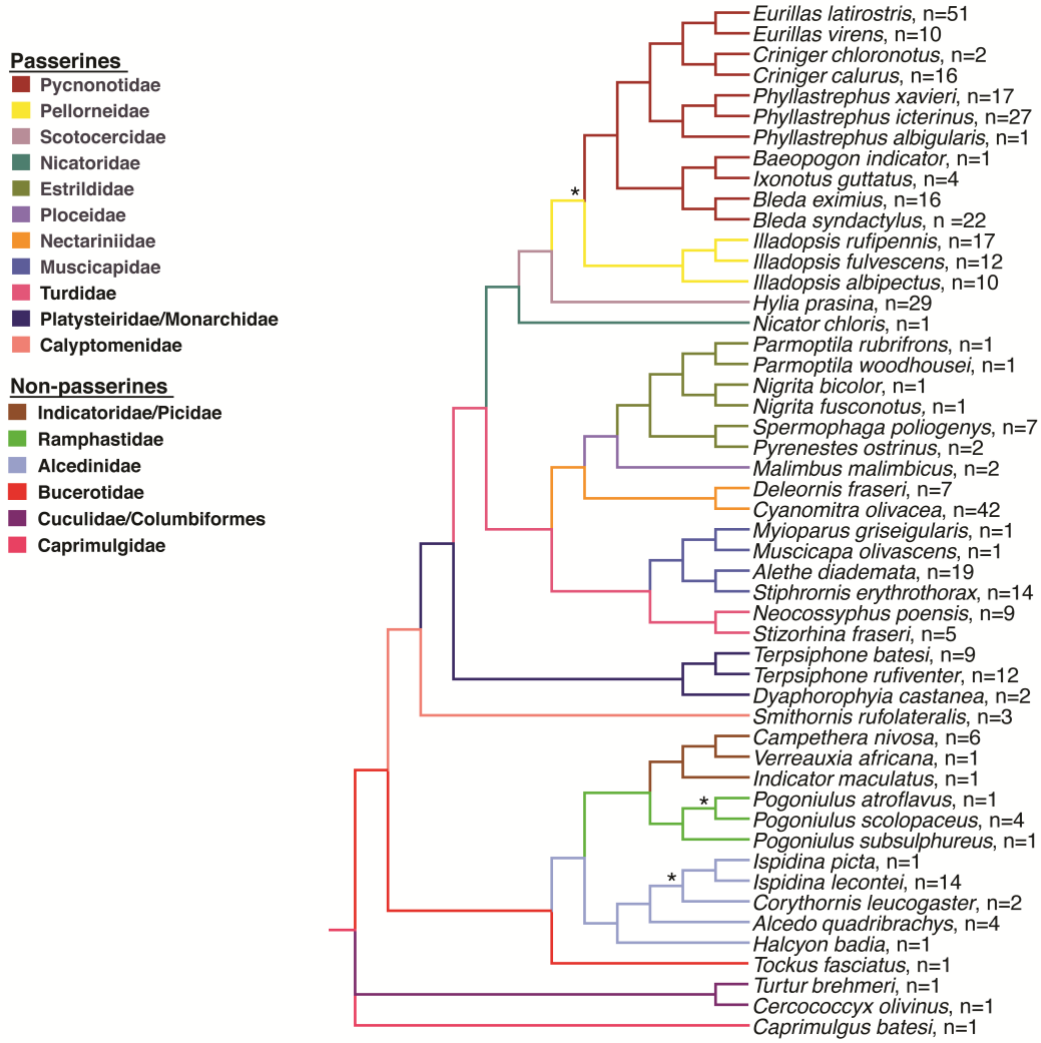

Figure 12. Avian consensus tree constructed using birdtree.org. Resolved from 9,000 distribution trees using the Hackett et al. (Hackett et al., 2008) as the backbone phylogeny. DRC avian sampling colored by clades which included order or family for descriptive purposes (Posterior probability ≤ 95 notated with an *).

in 42 of 421 individuals (9.98%), and *Plasmodium* in 124 of 421 individuals (29.5%). We recovered multiple infections, meaning two or more haemosporidian lineages, from the same host individual for 44 individuals.

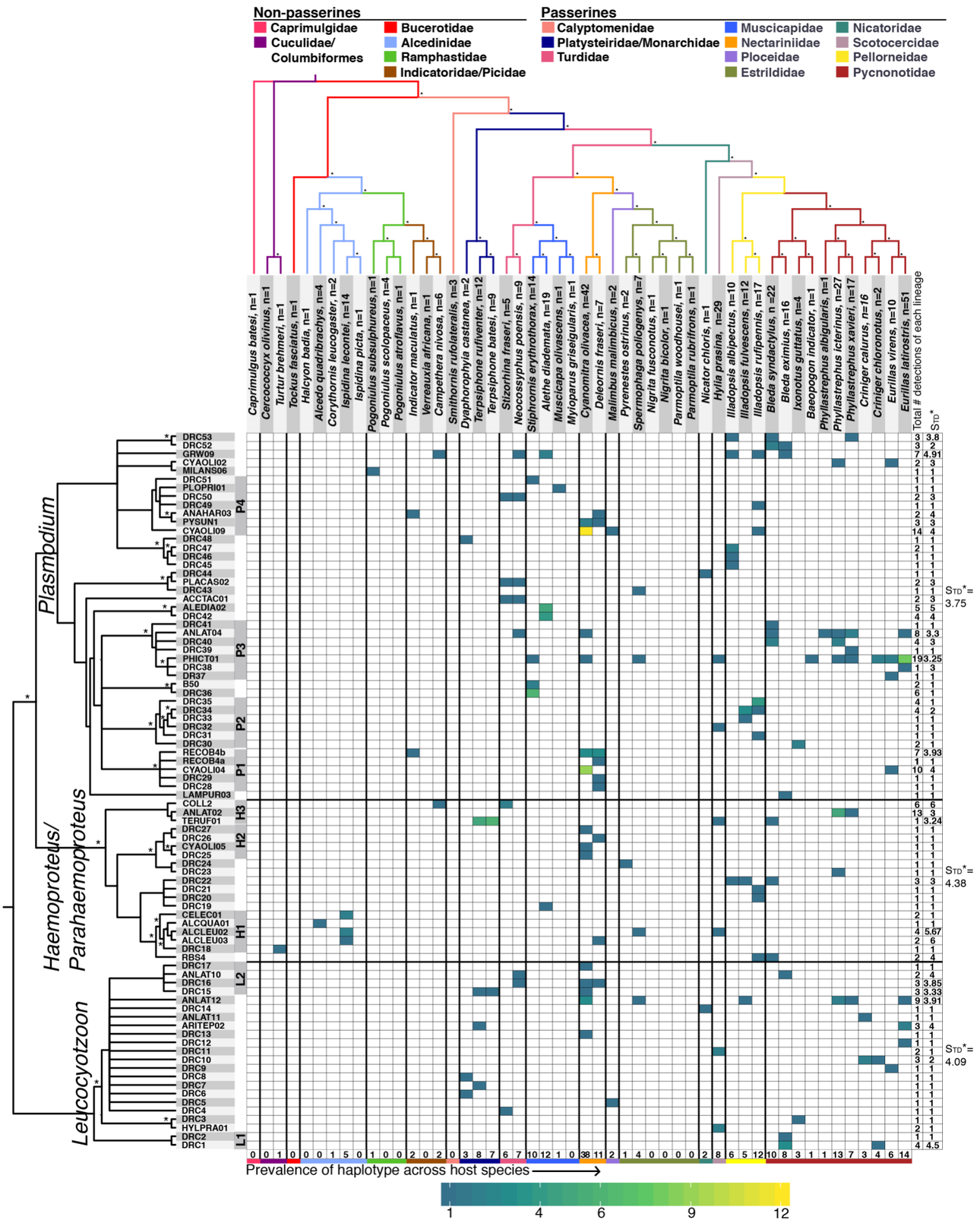

Figure 13. DRC avian host phylogeny (Birdtree.org, built from consensus of 10,000 trees, support ≤ 95 notated with an *) with each taxonomic grouping denoted by color group. The heat map shows prevalence of haemosporidian lineages (left, shown in phylogenetic order) as indicated by color gradient (1-12, zeroes noted with a white square) across avian host taxonomy. Prevalence total for each species sampled shown at bottom. On the right side is the prevalence of each lineage, the. STD* for each lineage, and the STD* average across each haemosporidian genus.

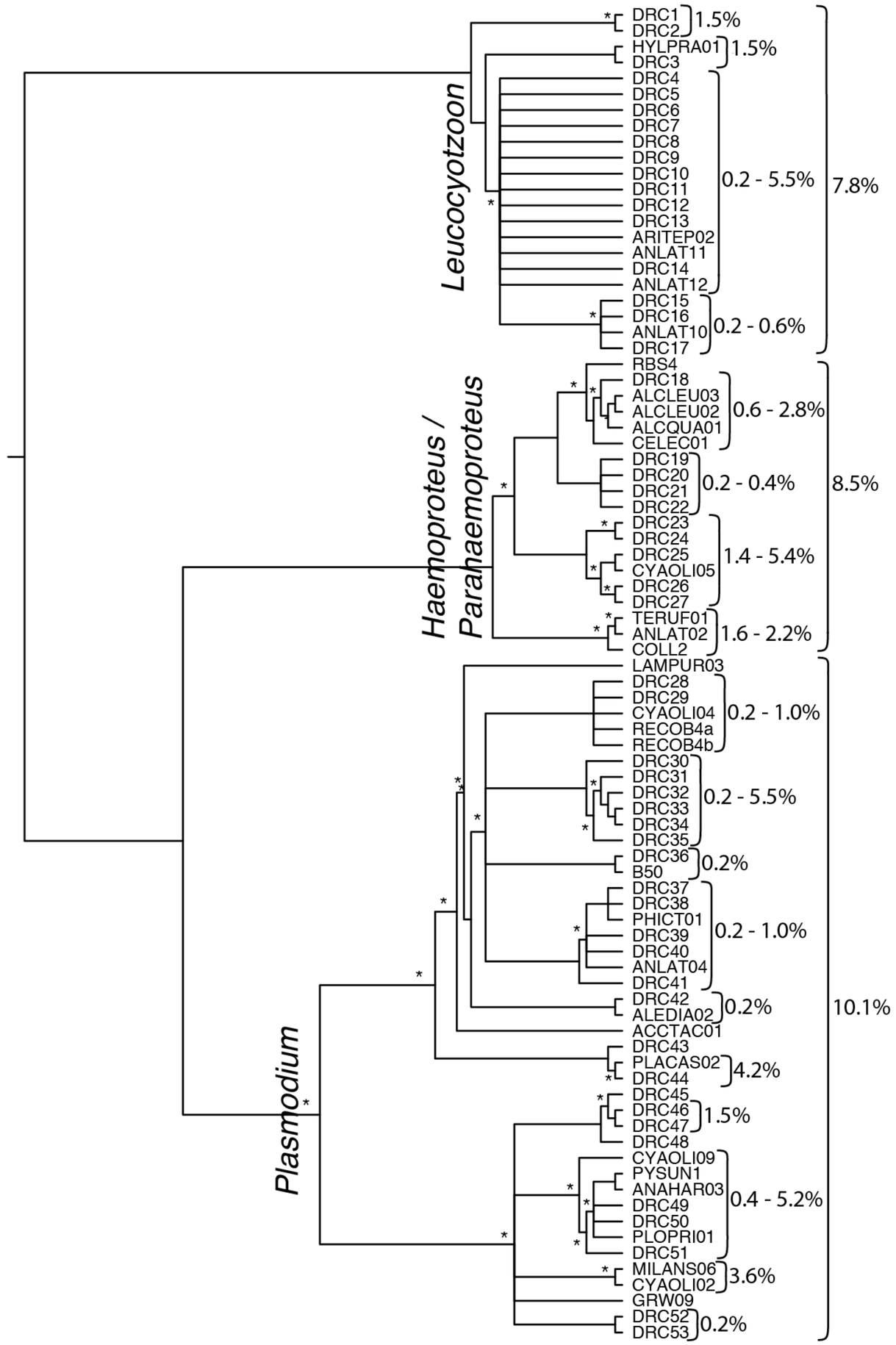

Figure 14. Cyt b haemosporidian phylogeny for unique haplotypes across DRC localities. Bayesian analysis with posterior probability ≥ 95 notated with an *. Pairwise sequence divergence for intra-generic clades (range of percent p-distance) and for each genus (maximum p-distance given).

Taxonomically, *Plasmodium* was recovered from all Passeriformes families except for Calyptomenidae (Figure 13; Table 8). *Haemoproteus* was recovered from all Passeriformes families except for Calyptomenidae, Ploceidae, and Scotocercidae. *Leucocytozoon* was recovered from all Passeriformes families except Calyptomenidae and Muscicapidae (Figure 13; Table 8). We recovered haemosporidian lineages from all non-Passeriformes families except for Camprimulgidae, Cuculidae, and Bucertotidae (Figure 13; Table 8). We recovered only *Haemoproteus* for the non-Passeriformes families Columbidae, Alcedinidae, and Picidae. We recovered *Plasmodium* from the non-Passeriformes families Ramphastidae, Indicatoridae, and Picidae. There were no *Leucocytozoon* recoveries in non-Passeriformes.

Pairwise genetic distance within each haemosporidian genus varied slightly, where *Haemoproteus* reached 8.5%, *Plasmodium* reached 10.1%, and *Leucocytozoon* reached 7.3% (Figures 13 and 14). Between group mean distance across genera was 10.4% for *Haemoproteus* and *Plasmodium*, 15.6% between *Haemoproteus* and *Leucocytozoon*, and 17.6% between *Plasmodium* and *Leucocytozoon*. Haemosporidian assigned clades (P1-P4, H1-H3, L1-L2) have a within clade pairwise divergence ranging between 0.6 - 5.5 % (Figures 13 and 14).

We recovered several haemosporidian clades which were highly associated with avian taxonomic clades (Figure 13). *Plasmodium* clade P1 (n = five lineages and 20 detections), which includes three previously detected lineages (MalAvi) and two novel lineages, was predominantly recovered from Nectariniidae with the exception of two recoveries from non-Nectariniidae lineages: once in *Indicator maculatus* (Indicatoridae)

and once in *Eurillas virens* (Pycnonotidae). These two exceptions correspond to MalAvi lineages CYAOLI04 (25 previous recoveries with 24 from Nectariniidae) and RECOB4 (31 previous recoveries with 27 from Nectariniidae), which collectively have been reported as recoveries in the Pycnonotidae, Estrildidae and Tyrannidae families. Our recoveries add a non-Passeriformes family and bring the overall total to just seven non-Nectariniidae recoveries for this *Plasmodium* clade, suggesting a preference for specificity to Nectariniidae.

Plasmodium clade P2 (n = six lineages and 17 detections) included only novel lineages and was recovered from members of Pellorneidae except for two lineages that were recovered from sister families Scotocercidae and Pycnonotidae (novel lineages P1-DDM1774B and BDM1671B) (Figure 13). *Plasmodium* clade P3 (n = seven lineages and 35 detections) was predominantly recovered from the family Pycnonotidae but was also recovered from five other Passeriformes families. Clade P3 contains five novel lineages and two previously detected two lineages, PHICT01 (three Nectariniidae recoveries and seven Pycnonotidae recoveries) and ANLAT04 (one Muscicapidae recovery and 12 Pycnonotidae recoveries). Overall, PHICT01 was the lineage with most detections (n = 19), and it was recovered from a total of five families (Figure 13). Based on MalAvi records, PHICT01 has been recovered in other Pycnonotidae and Nectariniidae species and has been recovered from three central African countries and Tanzania, indicating that this lineage is geographically widespread across the African tropical forest belt. While we recovered ANLAT04 in three families (Pycnonotidae, Turdidae, and Nectariniidae), MalAvi recoveries have been primarily in Pycnonotidae,

and geographically from central African countries. The exception is one Muscicapidae (*Ficedula hypoleuca*) recovery from the Russian Ural Mountains, indicating migratory transmission of this lineage. Overall, our ANLAT04 recoveries indicate a previously undetected level of host generalism in this lineage. All other lineages in clade P3 are novel.

For *Haemoproteus* clades, we see varied levels of host association. Clade H1 contains one novel lineages (DRC18) and four previously detected lineages (MalAvi lineages CELEC01, ALCQUA01 ALCLEU02, and ALCLEU03; Figure 13). Previous MalAvi associations for clade HI include Alcedinidae, Stenostiridae, Timalidae, and Estrildidae. Lineages are here primarily recovered from the family Alcedinidae and further recovered from Columbidae, Estrildidae, Nectariniidae, and Scotocercidae, representing new associations and an expanded host breadth for this haemosporidian clade.

In the *Haemoproteus* clade H2 we recovered one MalAvi lineage (CYAOLI05) and three novel lineages all from the family Nectariinidae (Figure 13). CYAOLI05 was here and previously only recovered in Nectariniidae. All novel H2 lineages are only recovered from Nectariniidae as well, indicating high host specificity.

For *Leucocytozoon,* Clade L1 is composed of two novel lineages (DRC1 and DRC2) which were recovered five times and solely from Pycnonotidae (Figure 13). Otherwise, for the clade L2 we recover three novel lineages and one previously recovered lineage, ANLAT10, which is here recovered from Turdidae and Pycnonotidae, while the four previous MalAvi recoveries include Pycnonotidae, Ploceidae, and

Alcedinidae. Other lineages recovered for *Leucocytozoon* resulted in no phylogenetic resolution for clades, while genetic distance remained low (< 5.6 %). Within these lineages is the most highly recovered *Leucocytozoon* lineage ANLAT12 (n = nine detections) which has been previously detected from the families Muscicapidae, Pycnonotidae, and Turdidae (Figure 13). These previous ANLAT12 recoveries are from Ghana (n = one) and Germany (n = two) and include two European-African migrants; *Turdus philomelos* which does not extend past northern Africa for wintering and *Musciapa striata* whose wintering grounds extend throughout most of sub-Saharan Africa. We detected this lineage, ANLAT12, in four families Pycnonotidae, Nectariniidae, Pellorneidae, and Estrildidae, with the latter three representing new associations (Figure 13).

Given the high sampling density across certain species and families we were able to assess the relationships of haemosporidians to several avian host families. Our sampling for the family Pycnonotidae includes 11 species across six genera (*Eurillas* = 61, *Criniger* = 18, *Phyllastephus* = 52, *Baeopogon* =1, *Ixonotus* = 4, and *Bleda* = 28; Figure 15). We recovered 70 haemosporidian detections for this family representing all three avian haemosporidian genera. *Haemoproteus* was restricted to two genera, *Phyllastrephus* and *Bleda* (Figure 15). *Leucocytozoon* was found across all species except for *Baeopogon indicator*. *Plasmodium* was found and demonstrated shared haplotypes across all 11 Pycnonotidae species. We recovered 28 unique haplotypes (*Haemoproteus* = four, *Leucocytozoon* = 10, *Plasmodium* = 14), reflecting high haemosporidian diversity as haplotypes range from one to 30 mutational steps from the

next nearest haplotype in *Plasmodium*, up to 24 mutational steps from the next nearest haplotype in *Leucocytozoon*, and up to 20 mutational steps from the next nearest haplotype in *Haemoproteus*.

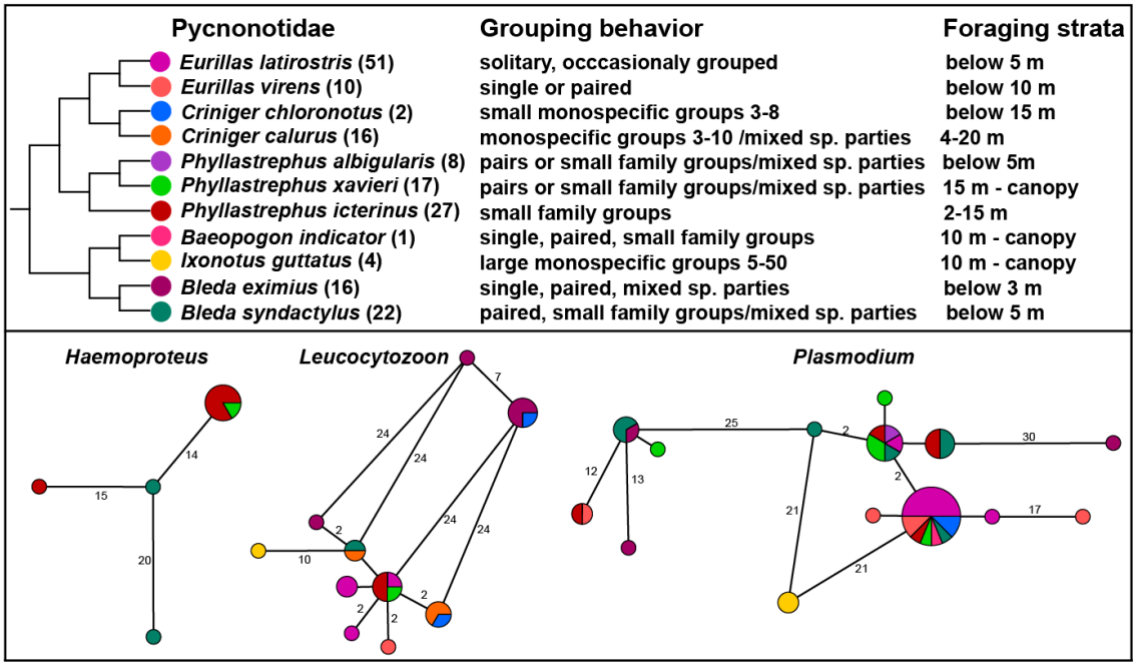

Figure 15. Haplotype network of haemosporidians detected in the avian family Pycnonotidae. Avian host phylogeny from Birdtree.org with description of grouping behavior and foraging strata for each species. Minimum-spanning networks for *Plasmodium*, *Leucocytozoon*, and *Haemoproteus* colored by host species associations.

Pycnonotidae species vary greatly in grouping behavior (Figure 15). Some species exhibit primarily solitary behaviors (e.g., *E. latirostris*) while some are found single or paired (e.g. *E. virens*). Other species form small family groups (e.g., all *Phyllasterephus* species, *C. choloronotus*) while some may occasionally join mixed species parties (e.g. all *Bleda* species, *C. calurus*, *P. albigularis*, *P. xavieri*). Finally, one

Pycnonotidae species in our sampling differs from others and is highly gregarious and found preferentially in large monospecific groups (*I. guttatus*) (Figure 15).

When examining the haplotype networks (Figure 15), for *Haemoproteus* we recovered three lineages restricted to a single host, and these range from 14 -20 mutational steps away from each other. These single host haplotypes were recovered in *B. syndactylus* and *P. icterinus*, with one shared haplotype among them. In *Leucocytozoon* the majority of haplotypes, eight of ten, are clustered together (two - ten mutational steps away from the next nearest neighbor). The two distant *Leucocytozoon* haplotypes were recovered in *B. eximius* and *C. chloronotus*. In *Plasmodium*, we recovered two frequently recovered haplotypes that were shared among nearly all Pycnonotid species. However, the remaining *Plasmodium* haplotypes are primarily restricted to single species and are often very distinct, with the most divergent haplotypes having 12 - 30 mutational steps to the next nearest haplotype. This is seen in haplotypes found in *B. eximius* (haplotypes at 13 and 30 mutational step distance) and *E. virens* (haplotypes at a 17-mutational step distance). All species of Pycnonotidae were recovered in at least one shared haplotype (shared lineage recovery with another Pycnonotidae species) except for *I. guttatus*. which is recovered three times for two novel lineages (DRC3 and DRC30, at 10 and 21 mutational steps from the next nearest neighbor, respectively) neither of which was recovered in any other species.

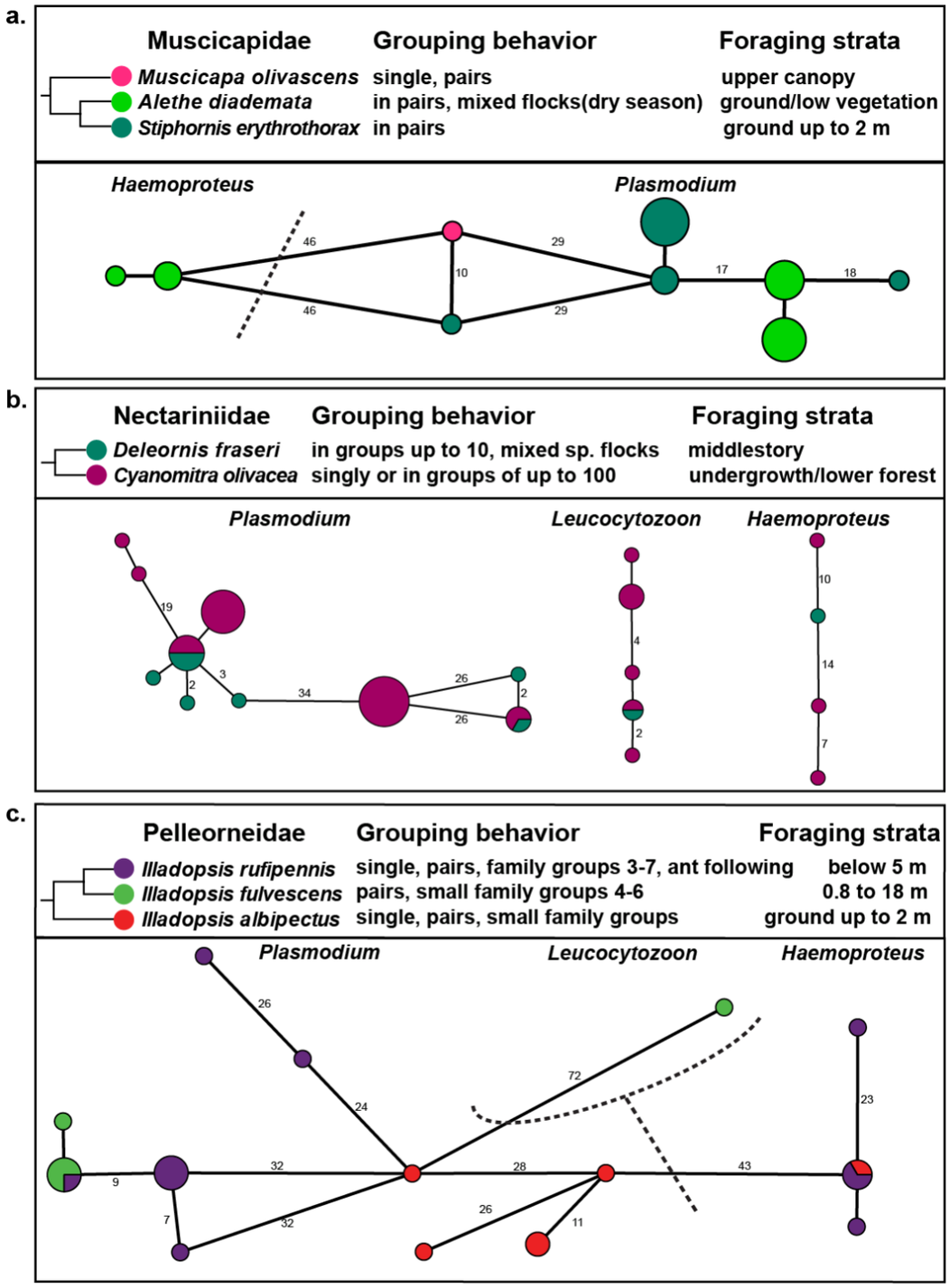

Figure 16. Haplotype network of haemosporidians detected in the avian families a) Muscicapidae b) Nectariniidae and c) Pellorneidae. Avian host phylogenies from Birdtree.org with description of grouping behavior and foraging strata for each species when information is available. Minimum-spanning networks for *Plasmodium, Leucocytozoon,* and *Haemoproteus* colored by host species associations. Haplotypes, colored for host species associations, are scaled for frequency of detection by size of circle. Mutation steps between haplotypes are indicated by the line and number or number of dashes.

We also had high sampling for species within the family Muscicapidae, which was represented here by three species (*Muscicapa olivascens* = 1, *Alethe diademata* = 19, and *Stiphonis erythrothorax* = 14 (to include the recently recognized *S. rudderi*; Voelker *et al.* 2017; Figure 16a). Previous detections for these species are limited, with six previous detections in both *A. diademata* and *S. erythrothorax*, and no previous detections in *M. olivascens*. We recovered nine unique haplotypes (*Haemoproteus* = two, *Plasmodium* = seven; Figure 16a), with no *Leucocytozoon* recoveries. There was no sharing of haplotypes among Muscicapidae species. Two of the *Plasmodium* haplotypes (MalAvi GRW09 and PHICT01) recovered here are host generalist parasites that have been recovered previously in multiple families, while all other seven lineages are novel and limited to Muscicapidae (Figure 13).

For the family Nectariniidae, our host sampling included two species, *Deleornis fraseri* (n = seven) and *Cyanomitra olivacea* (n = 42), with 49 recovered haemosporidian detections overall (Figure 16b). We recovered 19 unique haplotypes, of which nine haplotypes have been previously recorded in MalAvi, and the remaining ten represent novel detections. We recovered all three haemosporidian genera for both species of Nectariniidae (*Haemoproteus* = four, *Leucocytozoon* = five, *Plasmodium* = ten) with only three shared haplotypes between host species (two in *Plasmodium* and one in *Leucocytozoon*). The shared haplotypes include two previously recovered *Plasmodium* haplotypes (RECOB4 and PYSUN1) and one novel *Leucocytozoon* haplotype from clade L3 (Figure 16b). *Plasmodium* haplotypes were the most differentiated, ranging one to 34 steps from the next nearest haplotype.

Finally, for the avian family Pellorneidae we sampled three members of the genus *Illadopsis* (Figure 16c) with sampling across species ranging from ten to 17 (Figures 12 and 13). We recovered all three haemosporidian genera with 14 unique haplotypes (*Haemoproteus* = three, *Leucocytozoon* = one, *Plasmodium* = ten). Haplotypes within *Plasmodium* and *Haemoproteus* were highly differentiated, and ranged from one to 32 mutational steps apart for *Plasmodium* (with most being 24 steps or greater), and one to 23 mutational steps apart for *Haemoproteus*. We found two shared haplotypes between host species, with one from *Haemoproteus* (H1) and one from *Plasmodium* (P1).

Host specificity as measured by S_{TD}* (taxonomic distance of hosts weighted by prevalence) varied across lineages and genera (Figure 13). Many lineages were detected in a single host (overall 64% of 84 unique lineages; S_{TD}* = 1), while the remaining lineages resulted in S_{TD}* values ranging from two to six, where six was the maximum number of taxonomic steps. The host range was assessed for each haemosporidian genus (Figure 13), with *Haemoproteus* returning the highest average (S_{TD}* = 4.38), indicating a large host taxonomic range, followed by *Leucocytozoon* (S_{TD}* = 4.09), and the most host specific average S_{TD}* for a genus was *Plasmodium* (S_{TD}* = 3.75). When examining S_{TD}* across lineages, the most host generalist lineage for *Plasmodium* was GRW09 (S_{TD}* = 4.91), which here is found across five families and had an S_{TD}* of 4.91. Lineage RECOB4 was next (S_{TD}* = 3.8), found across 3 families. For *Haemoproteus* the most host generalist lineages were COLL2 (S_{TD}* = 6.0) and ALCLEU02 (S_{TD}* = 5.67) which were recovered in two and three families, respectively. When looking at host generalism

in the genus *Leucocytozoon* we found our most generalist lineage was novel lineage DRC1 (S_{TD}* = 4.5) followed by ANLAT10 (S_{TD}* = 4.0) recovered in one and two families, respectively.

We found that the addition of existing data (MalAvi) to our recovered data altered the previously known host specificity metric (Figure 17). As determined here, local data (our samples) S_{TD} was higher than global (our samples plus Malavi) S_{TD} in 11 of the 19 lineages assessed. Whereas, global S_{TD} was higher than local in four of the lineages and the remaining four lineages showed equal measures of host specificity (Figure 17). The variation of local and global S_{TD} was most extreme in *Haemoproteus* where four of five lineages demonstrated higher local than global S_{TD} measures (Figure 17).

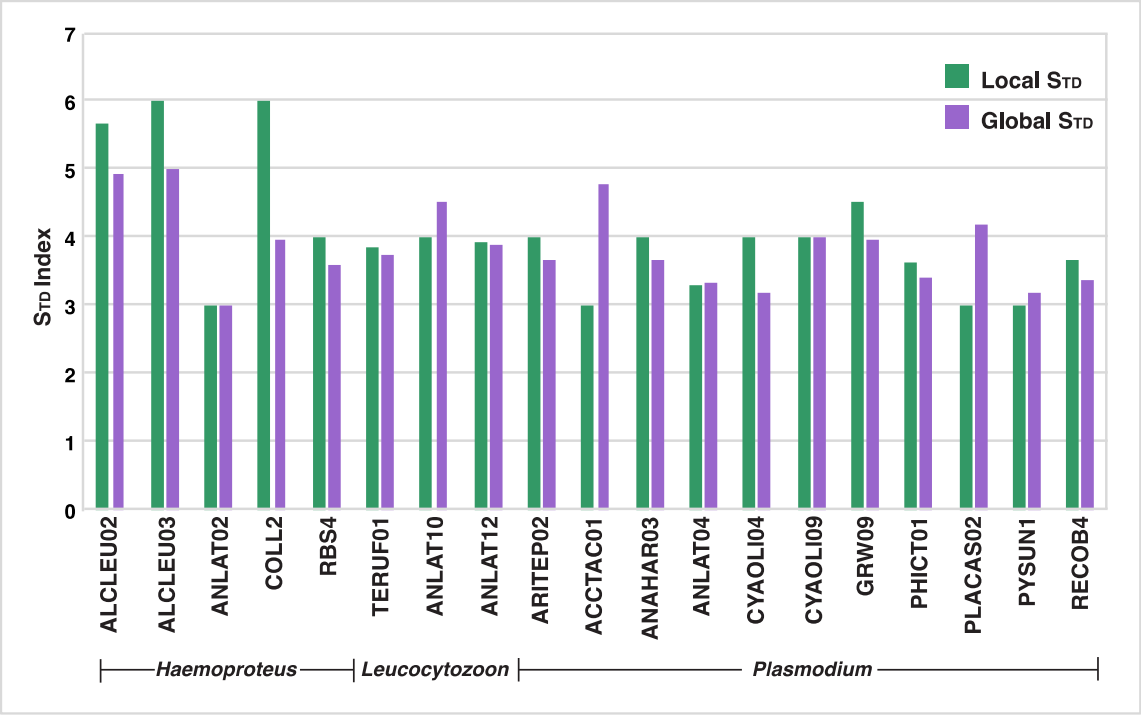

Figure 17. Host specificity measured by S_{TD} for each parasite lineage that was recovered in this study (local S_{TD}) and our recoveries with the addition of previously recorded MalAvi data and host associations (global S_{TD}). Parasite lineages recorded once are not included.

IV.4 Discussion

We report novel diversity recovered from the first assessment of haemosporidian infections in lowland tropical forest birds from the Democratic Republic of the Congo. Haemosporidian infection rates varied and were lower for *Haemoproteus* and *Plasmodium* and higher for *Leucocytozoon* than a previous community level analysis in central African lowland forest (Beadell *et al.* 2009). The variation in infection rate may be due to host densities, and transmission rates. Furth the proximity of our sampling sites to the Congo River. may impact vector distributions and densities, given the varying

requirement of standing and flowing water. Proximity to the river may increase black fly and therefore *Leucocytozoon* densities (Sutcliffe 1986; Ya'cob *et al.* 2016). Whereas, mosquitos are more variable in their constraints favoring a wide range of preferences for pooled water and a higher preference for habitats (i.e. upland swamps and forests vs open habitats) (Njabo *et al.* 2009, 2011) . Only a few studies have examined vectors and their associated haemosporidian infections in sub-Saharan Africa (Njabo *et al.* 2009, 2011; Njabo, Smith & Yohannes 2013; Ajamma *et al.* 2016). New association for genera vectoring avian haemosporidians are still being recovered (Njabo *et al.* 2009), with much of Africa still lacking vector studies.

We determined host associations for lineages and clades across all three haemosporidian genera. We found that host lineage constraint, as indicated by the STD* metric, demonstrated the highest host generalism values in *Haemoproteus*, especially in lineage COLL within clade H3 which had an S_{TD}* = 6.0, and lineages ALCLEU02 and ALCLEU03 within clade H1, which had S_{TD}* values of 5.67 and 6.0, respectively (Figure 13). We also found host specificity variation within clades, as seen in *Haemoproteus* clade H1, where S_{TD}* values range from 1- 6. We recovered this variation of S_{TD}* values across clades for *Plasmodium* (clades P1, P3 and P4) and *Leucocytozoon* (clades L1 and L2); however, clades did not discretely exhibit specificity or generalism. These patterns of association were seen across lineage clades and not across genera, supporting the previous findings for lineage level differentiation (Medeiros *et al.* 2014). Host specificity indices may clearly be affected by sampling bias, which is likely not only from our sampling methodology (see Methods) but also

from the fact that Africa has been poorly sampled for haemosporidians generally. Nonetheless, this is an important metric to examine as sampling of lineages increases.

Despite probable sampling bias across avian taxonomy a number of species are demonstrating interesting patterns. The patterns in the haemosporidians found in Muscicapidae demonstrate high species specificity, as no haplotypes were shared among the three species in our sampling. Of the 13 haplotypes/lineages recovered for Muscicapidae (3 species), zero haplotypes where shared among them. This result may be indicative of their grouping behavior where species are found single or in pairs and not in mixed groups, thereby reducing the potential encounter rate for shared parasites (Figure 16). If some haemosporidian lineages are indeed tending toward specificity at the host family level, reduction of host encounter rate with the same vectors might also be influenced by foraging strata; one of the African endemic muscicapids in our sampling forages in a distinctly different stratum that do the other two (Figure 16a). However, the remaining two haplotypes/lineages that we recovered from Muscicapidae (GRW09 and PHICT01, both *Plasmodium*), are widely shared among other avian families. Further analysis and modeling of life history traits may be informative in modeling haemosporidian distributions, traits such as foraging strata can be used as a proxy for vectors given foraging height of vectors themselves is not yet broadly available across vector diversity. This was true within the diverse family Pycnonotidae, whose species exhibit highly variable traits for grouping behaviors and foraging strata; yet, the diversity of lineages and host sharing was markedly high for *Plasmodium* and *Leucocytozoon* and limited for *Haemoproteus*.

Given the high number of detections (n = 49), the low number of shared haplotypes may be associated with the non-overlapping foraging strata preference of *D. fraseri* (middlestory foraging) and *C. olivacea* (undergrowth/lower forest foraging). The shared haplotypes recovered in the family Pellorneidae may be due to the overlap in foraging strata across these species, as they are all found between 0.8 to two meters (del Hoyo *et al.* 2017).

While we increased the avian taxonomic sampling for haemosporidian avian hosts in central Africa sampling in this region remains sampling poor, with taxonomically diverse sampling existing for only Gabon and Cameroon, with sampling across avian diversity low in Equatorial Guinea (Beadell *et al.* 2009; Bensch *et al.* 2009). We need increased sampling to confirm low haemosporidian competence of certain avian species and families. For example, we recovered no haemosporidians for the family Calyptomenidae, the African and Green Broadbills, which included sampling for three individuals of *Smithornis rufolateralis*. There is only one previous recovery of a haemosporidian for any member of Calyptomendidae, a *Leucoyctozoon* (MalAvi lineage AFR175) recovery in *Smithornis capensis*, perhaps explained by the broader habitat preferences of *S. capenis* as compared to other members of its family. Furthermore, an increase in sampling will elucidate connectivity and transmission of the high number of singleton lineages (recovered in one host species) detected. Our recoveries included a number of lineages that were recovered once or twice in MalAvi: ALCLEU02, ANAHAR03, ARITEP02, and TERUF01. The number of novel lineages detected in this study demonstrate the importance of not only increasing sampling

geographically but the importance of broadly assessing avian communities (Light *et al.* 2016; Walther *et al.* 2016; Marroquin-flores *et al.* 2017) to determine patterns of parasite composition along with host interactions (Hamilton & Zuk 1982; Lutz *et al.* 2015; Light *et al.* 2016).

CHAPTER V

CONCLUSIONS

Climate change is already affecting avian host and vector dynamics, we are beginning to see the impacts on disease dynamics and distributions. Studies suggest that these changes will not only shift or broaden parasite distributions but will also result in increased prevalence of parasites with global warming (Garamszegi 2011; Lapointe, Atkinson & Samuel 2012; Loiseau *et al.* 2013). Malaria has already been found in regions and taxa previously believed to be malaria free, such as Alaska and the Galapagos Islands (Levin *et al.* 2009; Loiseau *et al.* 2012a). We have also begun to see evidence of infections and mortality in species previously free of haemosporidian infections (Martinsen *et al.* 2017).

Determining predictors of presence and colonization of avian malaria is necessary to elucidate the how changes in abiotic factors (temperature, precipitation, seasonality, etc.) will alter the distribution of avian malaria across the landscape. Further, will the distinct genera of avian malaria show differences in response to climate effects? This information will be important for averting emerging infectious diseases and conservation prioritization of affected species.

In this dissertation, I determined that variation is found across source materials (here blood and pectoral muscle) for detections of haemosporidian (Chapter I). However, this variation is not the same across all genera of haemosporidians. Instead the results for optimal source material varies across haemosporidian genus. These results highlight the

variation in haemosporidian life cycle and the location of the life stages within the vertebrate host. Suggesting that study design and selection of source materials be guided by the genus of interest in the study.

In further work, I conducted the first sampling of avian haemosporidians from Benin located in tropical West Africa (Chapter III). I recovered a 50.9% infection rate, which we attribute in part due to the sampling of multiple source materials (blood and pectoral muscle). The lineages recovered along with the associated sampling localities were used to characterize the complex multivariate environment which restricts and informs haemosporidians and their vector associations. The novel visualization analysis, adapted heat map analyses, I developed allowed for determination of clades of haemosporidians (for all three genera) which were associated with specific hosts and environments and conversely to determine clades which are not restricted by hosts association or environmental constraints. This resulted in a number of intra-generic distribution patterns which were detected. However, no strong patterns were recovered at the genus level, suggesting that haemosporidian diversity cannot be characterized at this level.

Lastly, I also conducted the first sampling of avian haemosporidians from the Democratic Republic of the Congo (Chapter IV) located in central Africa. I recovered a 42.4% infection rate. This included a large number of avian host species never sampled before resulting in a high percentage of novel lineages (60 out of 84, 71.4%) detected for this study. Environmental analysis was not conducted for this study due to the proximity of sampling localities and low variation among sites. However, visualization analysis for

the lineage detections and their prevalence associated with host and lineage phylogenies resulted again in intra-generic patterns for haemosporidian clades. Assessment of host species and families with high sampling revealed some patterns in lineage association with host species which were informed by host life history traits. Indicating the need for further analysis of more life history traits which can inform the relationship of parasites their associated vectors and the avian host. The number of novel lineages detected in these chapters (III and IV) demonstrate the importance of addressing regions not yet sampled (Figure 18). MalAvi (Version 2.2.3) currently has 666 lineages which have been recovered in sub-Saharan Africa. These studies add an additional 100 novel lineages (Figure 18). This sampling further increases the information of distributions and host associations for previously detected lineages. The number of lineages detected in these chapters (III and IV) demonstrate the importance of addressing regions not yet sampled by detecting novel lineages and by increasing the information on distributions and host associations for previously detected lineages.

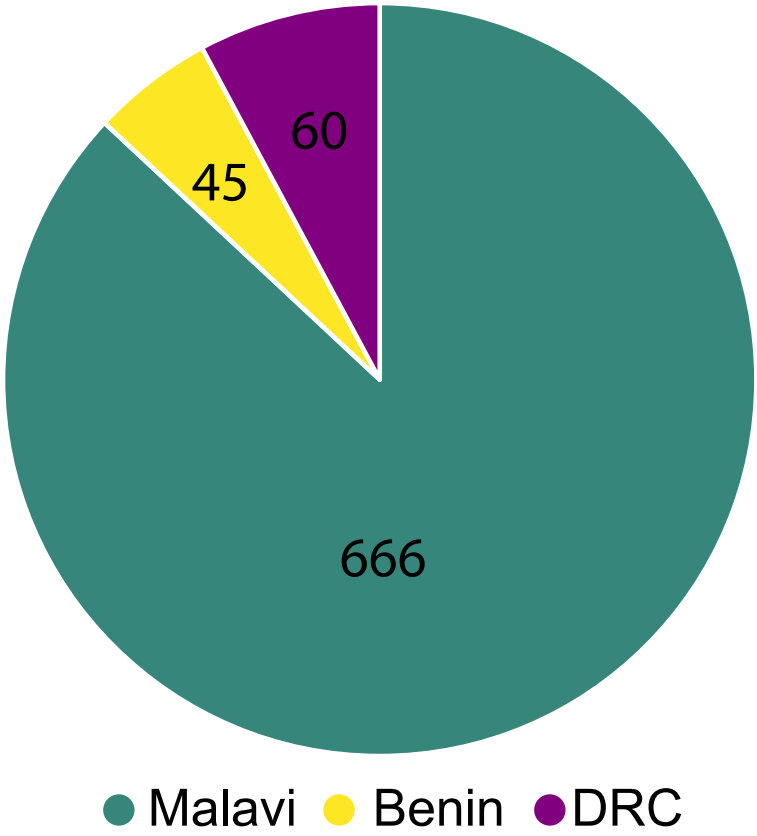

Figure 18. Pie chart of all lineages recovered from sub-Saharan Africa, including all previously detected MalAvi V2.2.3 lineages (teal), all novel lineages recovered in Benin (yellow), and all novel lineages recovered in the Democratic Republic of the Congo (DRC, purple).

Given these results there are two priorities for the future study of avian haemosporidian disease ecology.

1) We must determine the extent of avian haemosporidian diversity. This means an increase in surveys in ecoregions not yet sampled, with currently many regions remaining not sampled.

2) Sampling must target the broad diversity of avian taxonomy.

These priorities are necessary for three primary reasons. First, to monitor shifts in distribution of parasites out of their normal range extent. We know this extent is restricted by vector suitability and climate constraints of the parasite for development. Second, in order to determine infections in previously unaffected species. Lastly, in order to develop predictive models, as used in Chapter IV, the number of diverse regions sampled needs to be increased. This increase in well sampled regions will help strengthen the model's predictive ability and will then allow these models to be applied to regions not yet sampled. Assessing host and bioclimatic variables at a larger geographic scale and across multiple ecoregions will help to better elucidate these process that are regulating the distribution of haemosporidians.

REFERENCES

Ajamma, Y. U., J. Villinger, D. Omondi, D. Salifu, T. O. Onchuru, L. Njoroge, A. W. T. Muigai, and D. K. Masiga. 2016. Composition and genetic diversity of mosquitoes (Diptera: Culicidae) on Islands and mainland shores of Kenya's lakes Victoria and baringo. Journal of Medical Entomology 53:1348–1363.

Altizer, S., R. S. Ostfeld, P. T. J. Johnson, S. Kutz, and C. D. Harvell. 2013. Climate change and infectious diseases: from evidence to a predictive framework. Science (New York, N.Y.) 341:514–9.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. Journal of Molecular Biology 215:403–10.

Asghar, M., D. Hasselquist, and S. Bensch. 2011. Are chronic avian haemosporidian infections costly in wild birds? Journal of Avian Biology 42:530–537.

Atkinson, C. T., and C. Van Riper III. 1991. Pathogenicity and epizootiology of avian haematozoa: *Plasmodium*, *Leucocytozoan*, and *Haemoproteus*. Pages 19–48 *in* J. E. Loye and M. Zuk, editors.Bird-Parasite Interactions: Ecology, evolution, and behaviour. Oxford University Press, Oxford.

Bandelt, H. J., P. Forster, and A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16:37–48.

Barter, R. L., and B. Yu. Superheat: Jan 2017. An R package for creating beautigul and extendable heatmaps for visualizing complex data. http://arxiv.org/abs/1512.01524

Beadell, J. S., R. Covas, C. Gebhard, F. Ishtiaq, M. Melo, B. K. Schmidt, S. L. Perkins,

G. R. Graves, and R. C. Fleischer. 2009. Host associations and evolutionary relationships of avian blood parasites from West Africa. International journal for parasitology 39:257–66.

Beadell, J. S., E. Gering, J. Austin, J. P. Dumbacher, M. a Peirce, T. K. Pratt, C. T. Atkinson, and R. C. Fleischer. 2004. Prevalence and differential host-specificity of two avian blood parasite genera in the Australo-Papuan region. Molecular ecology 13:3829–44.

Beadell, J. S., F. Ishtiaq, R. Covas, M. Melo, B. H. Warren, C. T. Atkinson, S. Bensch, G. R. Graves, Y. V Jhala, M. a Peirce, A. R. Rahmani, D. M. Fonseca, and R. C. Fleischer. 2006. Global phylogeographic limits of Hawaii's avian malaria. Proceedings. Biological sciences / The Royal Society 273:2935–44.

Bensch, S., O. Hellgren, and J. Pérez-Tris. 2009. MalAvi: a public database of malaria parasites and related haemosporidians in avian hosts based on mitochondrial cytochrome b lineages. Molecular ecology resources 9:1353–8.

Bensch, S., J. Pérez-Tris, J. Waldenström, and O. Hellgren. 2004. Linkage between nuclear and mitochondrial DNA sequences in avian malaria parasites: multiple cases of cryptic speciation? Evolution; international journal of organic evolution 58:1617–21.

Bensch, S., M. Stjernman, D. Hasselquist, O. Ostman, B. Hansson, H. Westerdahl, and R. T. Pinheiro. 2000. Host specificity in avian blood parasites: a study of *Plasmodium* and *Haemoproteus* mitochondrial DNA amplified from birds. Proceedings. Biological sciences / The Royal Society 267:1583–9.

Bertram, M. R., S. A. Hamer, B. K. Hartup, K. F. Snowden, M. C. Medeiros, D. C. Outlaw, and G. L. Hamer. 2017. Molecular Phylogenetics and Evolution A novel Haemosporida clade at the rank of genus in North American cranes (Aves : Gruiformes). Molecular Phylogenetics and Evolution 109:73–79.

Brooks, T., A. Balmford, N. Burgess, J. O. N. Fjeldså, L. A. Hansen, J. Moore, C. Rahbek, and P. Williams. 2001. Toward a Blueprint for Conservation in Africa 51:613–624.

Browning, S., and B. Browning. 2011. Haplotype phasing: existing methods and new developments. Nature Reviews Genetics 12:703–714.

Calderaro, A., G. Piccolo, C. Gorrini, S. Rossi, S. Montecchini, M. L. Dell'Anna, F. De Conto, M. C. Medici, C. Chezzi, and M. C. Arcangeletti. 2013. Accurate identification of the six human *Plasmodium spp*. causing imported malaria, including *Plasmodium ovale wallikeri* and *Plasmodium knowlesi*. Malaria Journal 12:1–6.

Caminade, C., S. Kovats, J. Rocklov, A. M. Tompkins, A. P. Morse, F. J. Colón-González, H. Stenlund, P. Martens, and S. J. Lloyd. 2014. Impact of climate change on global malaria distribution. Proceedings of the National Academy of Sciences 111:3286–3291.

Clark, N. J., S. M. Clegg, and M. R. Lima. 2014. A review of global diversity in avian haemosporidians (*Plasmodium* and *Haemoproteus*: Haemosporida): new insights from molecular data. International Journal for Parasitology. 44: 329-338.

Comstedt, P., S. Bergström, B. Olsen, U. Garpmo, L. Marjavaara, H. Mejlon, A. G.

Barbour, and J. Bunikis. 2006. Migratory passerine birds as reservoirs of Lyme borreliosis in Europe. Emerging Infectious Diseases 12:1087–1095.

Coon, C. A. C., L. Garcia-Longoria, L. B. Martin, S. Magallanes, F. de Lope, and A. Marzal. 2016. Malaria infection negatively affects feather growth rate in the house sparrow *Passer domesticus*. Journal of Avian Biology 47:779–787.

Cox, F. E. 2010. History of the discovery of the malaria parasites and their vectors. Parasites and Vectors 3:5.

Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature methods 9:772.

Dickinson, E. C., and J. V. Remsen, editors. 2013. The Howard and Moore Complete Checklist of the birds of the world. 4th edition. Aves Press, Eastbourne, UK.

Dinerstein, E., D. Olson, A. Joshi, C. Vynne, N. D. Burgess, E. Wikramanayake, N. Hahn, S. Palminteri, P. Hedao, R. Noss, M. Hansen, H. Locke, E. C. Ellis, B. Jones, C. V. Barber, R. Hayes, C. Kormos, V. Martin, E. Crist, W. Sechrest, L. Price, J. E. M. Baillie, D. Weeden, K. Suckling, C. Davis, N. Sizer, R. Moore, D. Thau, T. Birch, P. Potapov, S. Turubanova, A. Tyukavina, N. De Souza, L. Pintea, J. C. Brito, O. A. Llewellyn, A. G. Miller, A. Patzelt, S. A. Ghazanfar, J. Timberlake, H. Klöser, Y. Shennan-Farpón, R. Kindt, J. P. B. Lillesø, P. Van Breugel, L. Graudal, M. Voge, K. F. Al-Shammari, and M. Saleem. 2017. An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm. BioScience 67:534–545.

Drovetski, S. V, S. a Aghayan, V. a Mata, R. J. Lopes, N. a Mode, J. a Harvey, and G. Voelker. 2014. Does the niche breadth or trade-off hypothesis explain the

abundance-occupancy relationship in avian Haemosporidia? Molecular Ecology 23:3322–3329.

Fick, S. E., and R. J. Hijmans. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37:4302–4315.

Fishpool, L., and J. Tobias. 2017. Common Bulbul (Pycnonotus barbatus). Page *in* Handbook of the Birds of the World Alive. J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana, editors. Lynx Edicions, Barcelona.

Gage, K. L., T. R. Burkot, R. J. Eisen, and E. B. Hayes. 2008. Climate and Vectorborne Diseases. American Journal of Preventive Medicine 35:436–450.

Gager, A. B., J. Del Rosario Loaiza, D. C. Dearborn, and E. Bermingham. 2008. Do mosquitoes filter the access of *Plasmodium* cytochrome b lineages to an avian host? Molecular Ecology 17:2552–2561.

Garamszegi, L. Z. 2011. Climate change increases the risk of malaria in birds. Global Change Biology 17:1751–1759.

Garcia-Longoria, L., L. Z. Garamszegi, and a. P. Moller. 2014. Host escape behavior and blood parasite infections in birds. Behavioral Ecology 25: 890-900.

Garcia-Longoria, L., O. Hellgren, S. Bensch, F. De Lope, and A. Marzal. 2015. Detecting transmission areas of malaria parasites in a migratory bird species. Parasitology 142:1215–1220.

Garnham, P. C. C. 1966. Malaria Parasites and Other Haemosporidia. Page (Blackwell Scientific Publications, Ed.). Oxford.

González, C., O. Wang, S. E. Strutz, C. González-Salazar, V. Sánchez-Cordero, and S.

Sarkar. 2010. Climate Change and Risk of Leishmaniasis in North America: Predictions from Ecological Niche Models of Vector and Reservoir Species. PLoS Neglected Tropical Diseases 4:e585.

Guindon, S., O. Gascuel, and B. Rannala. 2003. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. Systematic Biology 52:696–704.

Hackett, S. J., R. T. Kimball, S. Reddy, R. C. K. Bowie, E. L. Braun, M. J. Braun, J. L. Chojnowski, W. A. Cox, K.-L. Han, J. Harshman, C. J. Huddleston, B. D. Marks, K. J. Miglia, W. S. Moore, F. H. Sheldon, D. W. Steadman, C. C. Witt, and T. Yuri. 2008. A Phylogenomic Study of Birds Reveals Their Evolutionary History. Science 320:1763–1768.

Hamilton, W. D., and M. Zuk. 1982. Heritable true fitness and bright birds: A role for parasites? Science 218:384–387.

Harvey, J. A., and G. Voelker. 2017. Avian haemosporidian detection across source materials: prevalence and genetic diversity. Parasitology Research 116:3361–3371.

Hellgren, O., J. Pérez-Tris, and S. Bensch. 2009. A jack-of-all-trades and still a master of some: prevalence and host range in avian malaria and related blood parasites. Ecology 90:2840–9.

Hellgren, O., J. Waldenström, J. Peréz-Tris, E. Szöll Osi, D. Hasselquist, A. Krizanauskiene, U. Ottosson, and S. Bensch. 2007. Detecting shifts of transmission areas in avian blood parasites: a phylogenetic approach. Molecular ecology 16:1281–90.

Van Hemert, C., J. M. Pearce, and C. M. Handel. 2014. Wildlife health in a rapidly changing North: Focus on avian disease. Frontiers in Ecology and the Environment 12:548–555.

del Hoyo, J., A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana, editors.. Handbook of the Birds of the World Alive. Lynx Edicions, Barcelona. http://www.hbw.com.

Illera, J. C., G. López, L. García-Padilla, and Á. Moreno. 2017. Factors governing the prevalence and richness of avian haemosporidian communities within and between temperate mountains. Plos One 12:e0184587.

Ishtiaq, F., J. S. Beadell, A. J. Baker, A. R. Rahmani, Y. V Jhala, and R. C. Fleischer. 2006. Prevalence and evolutionary relationships of haematozoan parasites in native versus introduced populations of common myna *Acridotheres tristis*. Proceedings. Biological sciences / The Royal Society 273:587–94.

Jarvi, S. I., M. E. Farias, D. A. Lapointe, M. Belcaid, and C. T. Atkinson. 2013. Next-generation sequencing reveals cryptic mtDNA diversity of *Plasmodium relictum* in the Hawaiian Islands. Parasitology 140:1741–1750.

Jetz, W., G. H. Thomas, J. B. Joy, K. Hartmann, and A. O. Mooers. 2012. The global diversity of birds in space and time. Nature 491:444–448.

Jetz, W., G. H. Thomas, J. B. Joy, D. W. Redding, K. Hartmann, and A. O. Mooers. 2014. Global distribution and conservation of evolutionary distinctness in birds. Current Biology 24:919–930.

de Jong, S. 1993. SIMPLS: An alternative approach to partial least squares regression. Chemometrics and Intelligent Laboratory Systems 18:251–263.

Joy, D. A. 2003. Early origin and recent expansion of *Plasmodium falciparum*. Science 300:318–321.

Keesing, F., R. D. Holt, and R. S. Ostfeld. 2006. Effects of species diversity on disease risk. Ecology Letters 9:485–498.

Kilpatrick, A. M., L. D. Kramer, M. J. Jones, P. P. Marra, and P. Daszak. 2006a. West Nile virus epidemics in North America are driven by shifts in mosquito feeding behavior. PLoS Biology 4:606–610.

Kilpatrick, A. M., D. A. LaPointe, C. T. Atkinson, B. L. Woodworth, J. K. Lease, M. E. Reiter, and K. Gross. 2006b. Effects of Chronic Avian malaria (*Plasmodium relictum*) infection on reproductive success of Hawaii Amakihi (*Hemignathus Virens*). The Auk 123:764.

Knowles, S. C. L., V. Palinauskas, and B. C. Sheldon. 2010. Chronic malaria infections increase family inequalities and reduce parental fitness: experimental evidence from a wild bird population. Journal of evolutionary biology 23:557–69.

Kumar, S., G. Stecher, and K. Tamura. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution 33:1870–1874.

Lamerton, J. F. 1965. A Key to the Genera of Hippoboscidae (Diptera: Pupipara), and to the Species of Hippobosca, in Africa. East African Agricultural and Forestry Journal 31:1–7.

Lanfear, R., P. B. Frandsen, A. M. Wright, T. Senfeld, and B. Calcott. 2017. Partitionfinder 2: New methods for selecting partitioned models of evolution for

molecular and morphological phylogenetic analyses. Molecular Biology and Evolution 34:772–773.

Lapointe, D. A., C.T. Atkinson, P.C. Banko, R. J. Camp, P. M. Gorresen, J.D. Jacobi, D.J. Jacobi, T.K. Thane and M.D. Samuel. 2013. Hawaiian Forest Birds : The past , present and future status of an endangered avifauna. USGS Staff Published Research. Paper 765.

LaPointe, D. A., M. L. Goff, and C. T. Atkinson. 2010. Thermal constraints to the sporogonic development and altitudinal distribution of avian malaria *Plasmodium relictum* in Hawai'i. Journal of Parasitology 96:318–324.

Lapointe, D. a, C. T. Atkinson, and M. D. Samuel. 2012. Ecology and conservation biology of avian malaria. Annals of the New York Academy of Sciences 1249:211–26.

Levin, I. I., D. C. Outlaw, F. H. Vargas, and P. G. Parker. 2009. Plasmodium blood parasite found in endangered Galapagos penguins (*Spheniscus mendiculu*s). Biological Conservation 142:3191–3195.

Librado, P., and J. Rozas. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452.

Light, J. E., C. E. Nessner, D. R. Gustafsson, S. R. Wise, and G. Voelker. 2016. Remarkable levels of avian louse (Insecta: Phthiraptera) diversity in the Congo Basin. Zoologica Scripta 45:538–551.

Linder, H. P., H. M. de Klerk, J. Born, N. D. Burgess, J. Fjeldså, and C. Rahbek. 2012. The partitioning of Africa: statistically defined biogeographical regions in sub-

Saharan Africa. Journal of Biogeography 39:1189–1205.

Loiseau, C., R. J. Harrigan, C. Bichet, R. Julliard, S. Garnier, A. Z. Lendvai, O. Chastel, and G. Sorci. 2013. Predictions of avian *Plasmodium* expansion under climate change. Scientific Reports 3:1126.

Loiseau, C., R. J. Harrigan, A. J. Cornel, S. L. Guers, M. Dodge, T. Marzec, J. S. Carlson, B. Seppi, and R. N. M. Sehgal. 2012a. First evidence and predictions of Plasmodium transmission in Alaskan bird populations. PloS one 7:e44729.

Loiseau, C., R. J. Harrigan, A. Robert, R. C. K. Bowie, H. a Thomassen, T. B. Smith, and R. N. M. Sehgal. 2012b. Host and habitat specialization of avian malaria in Africa. Molecular Ecology 21:431–41.

Lutz, H. L., W. M. Hochachka, J. I. Engel, J. a. Bell, V. V. Tkach, J. M. Bates, S. J. Hackett, and J. D. Weckstein. 2015. Parasite prevalence corresponds to host life history in a diverse assemblage of afrotropical birds and haemosporidian parasites. Plos One 10:e0121254.

Malmqvist, B., D. Strasevicius, O. Hellgren, P. H. Adler, and S. Bensch. 2004. Vertebrate host specificity of wild-caught blackflies revealed by mitochondrial DNA in blood. Proceedings of the Royal Society B: Biological Sciences 271:S152–S155.

Marroquin-flores, R. A., J. L. Williamson, A. N. Chavez, S. M. Bauernfeind, M. J. Baumann, C. R. Gadek, A. B. Johnson, J. M. McCullough, C. C. Witt, and L. N. Barrow. 2017. Diversity , abundance , and host relationships in the avian malaria community of New Mexico pine forests. PeerJ 5:e3700.

Martínez-De La Puente, J., J. Martínez, J. Rivero-De Aguilar, J. Herrero, and S. Merino. 2011. On the specificity of avian blood parasites: Revealing specific and generalist relationships between haemosporidians and biting midges. Molecular Ecology 20:3275–3287.

Martinsen, E. S., I. F. Sidor, S. Flint, J. Cooley, and M. A. Pokras. 2017. Documentation of Malaria parasite (*Plasmodium spp.*) infection and associated mortality in a Common Loon (*Gavia immer*). Journal of Wildlife Diseases 53:859–863.

Marzal, A., F. de Lope, C. Navarro, and A. P. Møller. 2005. Malarial parasites decrease reproductive success: an experimental study in a passerine bird. Oecologia 142:541–5.

Medeiros, M. C. I., V. a Ellis, and R. E. Ricklefs. 2014. Specialized avian Haemosporida trade reduced host breadth for increased prevalence. Journal of Evolutionary Biology 27:2520–2528.

Meillon, B. de, and W. W. Wirth. 1991. The genera and subgenera (excluding Culicoides) of the Afrotropical Biting Midges (Diptera : Ceratopogonidae). Annals of the Natal Museum 32:27–147.

Meiswinkel, R., G. J. Venter, and E. M. Nevill. 2004. Vectors: *Culicoides spp*. Infectious Diseases of Livestock 1:93–136.

Mellor, P. S., J. Boorman, and M. Baylis. 2000. *Culicoides* Biting Midges: Their role as arbovirus vectors. Annual Review of Entomology 45:307–340.

Merino, S., J. Moreno, J. J. Sanz, and E. Arriero. 2000. Are avian blood parasites pathogenic in the wild? A medication experiment in blue tits (*Parus caeruleus*).

Proceedings. Biological sciences / The Royal Society 267:2507–2510.

Mevik, B.-H., and R. Wehrens. 2007. The pls Package: Principle component and partial least squares regression in R. Journal of Statistical Software 18:1–24.

Mittermeier, R. a., N. Myers, J. B. Thomsen, G. a. B. Fonseca, and S. Olivieri. 1998. Biodiversity hotspots and major tropical wilderness areas : approaches to setting conservation priorities. Conservation Biology 12:516–520.

Njabo, K. Y., A. J. Cornel, R. N. M. Sehgal, C. Loiseau, W. Buermann, R. J. Harrigan, J. Pollinger, G. Valkiūnas, and T. B. Smith. 2009. Coquillettidia (Culicidae, Diptera) mosquitoes are natural vectors of avian malaria in Africa. Malaria Journal 8:193.

Njabo, K. Y., A J. Cornel, C. Bonneaud, E. Toffelmier, R. N. M. Sehgal, G. Valkiūnas, A F. Russell, and T. B. Smith. 2011. Nonspecific patterns of vector, host and avian malaria parasite associations in a central African rainforest. Molecular Ecology 20:1049–61.

Njabo, K. Y., T. B. Smith, and E. Yohannes. 2013. Feeding habits of culicine mosquitoes in the Cameroon lowland forests based on stable isotopes and blood meal analyses. Journal of Parasitology and Vector Biology 5:6–12.

Okanga, S., G. S. Cumming, P. a. R. Hockey, L. Nupen, and J. L. Peters. 2014. Host specificity and co-speciation in avian haemosporidia in the Western Cape, South Africa. PLoS ONE 9:e86382.

Olsson-pons, S., N. J. Clark, F. Ishtiaq, and S. M. Clegg. 2015. Differences in host species relationships and biogeographic influences produce contrasting patterns of prevalence , community composition and genetic structure in two genera of avian

malaria parasites in southern Melanesia. Journal of Animal Ecology. 84: 985-998.

Outlaw, D. C., J. A. Harvey, S. V. Drovetski, and G. Voelker. 2017. Diversity and distribution of avian haemosporidians in sub-Saharan Africa: an inter-regional biogeographic overview. Parasitology 144:394–402.

Outlaw, D. C., and R. E. Ricklefs. 2014. Species limits in avian malaria parasites (Haemosporida): how to move forward in the molecular era. Parasitology 141:1223–32.

Poulin, R. 2007. Host Specificity *in* Evolutionary Ecology of Parasites. Second. Princeton University Press, New Jersey. Pages 41–69.

Poulin, R., and D. Mouillot. 2005. Combining phylogenetic and ecological information into a new index of host specificity. The Journal of Parasitology 91:511–514.

R Core Team. 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Page (L. Ranjard, F. Poly, and J. C. Lata, Eds.). Vienna, Austria.

Randolph, S. E., and D. J. Rogers. 2010. The arrival, establishment and spread of exotic diseases: patterns and predictions. Nature Reviews. Microbiology 8:361–371.

Rich, S. M., and F. J. Ayala. 2006. Evolutionary Origins of Human Malaria Parasites *in* Malaria: Genetic and Evolutionary Aspects. K. Dronamraju and P. Arese, editors. Springer. Pages 125–143.

Ricklefs, R. E. 2010. Evolutionary diversification, coevolution between populations and their antagonists, and the filling of niche space. Proceedings of the National Academy of Sciences 107:1265–1272.

Ricklefs, R. E., and S. M. Fallon. 2002. Diversification and host switching in avian malaria parasites. Proceedings. Biological sciences / The Royal Society 269:885–92.

Ricklefs, R. E., S. M. Fallon, S. C. Latta, B. L. Swanson, and E. Bermingham. 2005. Migrants and their parasites: A bridge between two worlds. *in* Birds of Two Worlds; the ecology and evolution of migration. R. Greenberg and P. P. Marra, editors. Johns Hopkins University Press, Baltimore. Pages 210–221.

Ricklefs, R. E., D. C. Outlaw, M. Svensson-Coelho, M. C. I. Medeiros, V. a Ellis, and S. Latta. 2014. Species formation by host shifting in avian malaria parasites. Proceedings of the National Academy of Sciences of the United States of America, 111:14816-14821.

van Riper III, C., S. G. van Riper, M. L. Goff, and M. Laird. 1986. The epizootiology and ecological significance of malaria in Hawaiian land birds. Ecological Monographs 56:327.

Santiago-Alarcon, D., V. Palinauskas, and H. M. Schaefer. 2012. Diptera vectors of avian haemosporidian parasites: Untangling parasite life cycles and their taxonomy. Biological Reviews 87:928–964.

Sehgal, R. N. M., W. Buermann, R. J. Harrigan, C. Bonneaud, C. Loiseau, a Chasar, I. Sepil, G. Valkiūnas, T. Iezhova, S. Saatchi, and T. B. Smith. 2011. Spatially explicit predictions of blood parasites in a widely distributed African rainforest bird. Proceedings. Biological Sciences / The Royal Society 278:1025–33.

Smith, K. F., M. Goldberg, S. Rosenthal, L. Carlson, J. Chen, C. Chen, and S.

Ramachandran. 2014. Global rise in human infectious disease outbreaks. Journal of The Royal Society Interface 11:20140950–20140950.

Stephens, M., N. J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. The American Journal of Human Genetics 68:978–989.

Sutcliffe, J. F. 1986. Black fly host location: a review. Canadian Journal of Zoology. 64:1041–1053.

Szymanski, M. M., and I. J. Lovette. 2005. High lineage diversity and host sharing of malarial parasites in a local avian assemblage. The Journal of Parasitology 91:768–774.

Takano, O. M., P. S. Mitchell, D. R. Gustafsson, A. Adite, G. Voelker, and J. E. Light. 2017. An assessment of host associations, geographic distributions, and genetic diversity of avian chewing lice (Insecta: Phthiraptera) from Benin. Journal of Parasitology 103:152–160.

Title, P. O., and J. B. Bemmels. 2017. ENVIREM: An expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. Ecography 40:1–16.

Valkiunas, G. 2005. Avian Malaria Parasites and Other Haemosporidia. CRC Press, Boca Raton.

Voelker, G., M. Tobler, H. L. Prestridge, E. Duijm, D. Groenenberg, M. R. Hutchinson, A. D. Martin, A. Nieman, C. S. Roselaar, and J. W. Huntley. 2017. Three new species of *Stiphrornis* (Aves: Muscicapidae) from the Afro-tropics, with a

molecular phylogenetic assessment of the genus. Systematics and Biodiversity 15:87–104.

Walther, E. L., J. S. Carlson, A. Cornel, B. K. Morris, and R. N. M. Sehgal. 2016. First molecular study of prevalence and diversity of avian haemosporidia in a central California songbird community. Journal of Ornithology 157:549–564.

Ward, S. 1992. Assessing functional explanations of host-specificity. The American Naturalist 139:883–891.

Wood, M. J., C. L. Cosgrove, T. a Wilkin, S. C. L. Knowles, K. P. Day, and B. C. Sheldon. 2007. Within-population variation in prevalence and lineage distribution of avian malaria in blue tits, *Cyanistes caeruleus*. Molecular Ecology 16:3263–73.

World Malaria Report 2017. 2017. World Heatlh Organization. Geneva.

Ya'cob, Z., H. Takaoka, P. Pramual, V. L. Low, and M. Sofian-Azirun. 2016. Distribution pattern of black fly (Diptera: Simuliidae) assemblages along an altitudinal gradient in Peninsular Malaysia. Parasites and Vectors 9:219.